

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: February 25, 2003, 14:51:36 ; Search time 9045 Seconds  
(without alignments)  
11666.863 Million cell updates/sec  
Title: US-09-900-237-29  
Perfect score: 3626  
Sequence: 1 gcacgaggaacccgcctcca.....tcacattttggaggagatttt 3626  
Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_em: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_or: \*  
21: em\_ov: \*  
22: em\_pat: \*  
23: em\_ph: \*  
24: em\_pi: \*  
25: em\_ro: \*  
26: em\_sts: \*  
27: em\_un: \*  
28: em\_vi: \*  
29: em\_vt: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rtd: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrt: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2595	71.6	3795	8	AF200533	AF200533 Zea mays
2	2595	71.6	3799	6	AX338680	AX338680 Sequence
3	2572.2	70.9	3745	8	AF200528	AF200528 Zea mays
4	2325.2	64.1	3676	8	AF200529	AF200529 Zea mays
5	1699.8	46.9	3723	8	AF150630	AF150630 Gossypium
6	1675.2	46.2	3682	8	AF022174	AF022174 Arabidops
7	1673.6	46.2	3614	6	AX030946	AX030946 Sequence
8	1673.6	46.2	3614	6	AX030960	AX030960 Sequence
9	1408.4	38.8	3725	8	AF200526	AF200526 Zea mays
10	1373	37.9	3752	8	AF200525	AF200525 Zea mays
11	1343.4	37.0	3603	6	AX030942	AX030942 Sequence
12	1343.4	37.0	3603	8	AX030956	AX030956 Sequence
13	1341.8	37.0	3673	6	AX030948	AX030948 Sequence
14	1341.8	37.0	3673	8	AX030962	AX030962 Sequence
15	1340.2	37.0	3812	8	AF200532	AF200532 Zea mays
16	1331.4	36.7	2828	8	AF200527	AF200527 Zea mays
17	1331.4	36.7	2830	6	AX338676	AX338676 Sequence
18	1316.8	36.3	3538	8	AF200530	AF200530 Zea mays
19	1271	35.1	3968	8	AF200531	AF200531 Zea mays
20	1244.8	34.3	3847	8	AF304374	AF304374 Nicotiana
21	1221	33.7	3081	8	AF088917	AF088917 Arabidops
22	1215.6	33.5	3186	8	GHU58283	U58283 Gossypium h
23	1215.6	33.5	3328	6	AR164434	AR164434 Sequence
24	1192.6	32.9	3771	8	AF027173	AF027173 Arabidops
25	1177	32.5	3444	8	AF062485	AF062485 Arabidops
26	1176.6	32.4	3828	6	AX030944	AX030944 Sequence
27	1176.6	32.4	3828	8	AX030958	AX030958 Sequence
28	1171	32.3	87793	2	AP003837	AP003837 Oryza sat
29	1171	32.3	120868	2	AP003748	AP003748 Oryza sat
30	1166.4	32.2	2958	8	AF267742	AF267742 Arabidops
31	1164.2	32.1	3395	8	AF081534	AF081534 Populus x
32	1158.4	31.9	3109	8	AY099636	AY099636 Arabidops
33	1134.2	31.3	3145	8	AF323039	AF323039 Zinnia el
34	1133.6	31.3	3232	8	AF072131	AF072131 Populus t
35	984.2	27.1	2248	6	AX030938	AX030938 Sequence
36	984.2	27.1	2248	8	AX030952	AX030952 Sequence
37	918.2	25.3	2229	8	GHU58284	U58284 Gossypium h
c	746.4	20.6	1529	8	TAU48693	TAU48693 Triticum ae
39	684	18.9	165918	2	AP005420	AP005420 Oryza sat
40	684	18.9	188279	2	AP005579	AP005579 Oryza sat
41	666.2	18.4	7234	6	AX048245	AX048245 Sequence
42	666.2	18.4	7234	8	AF091713	AF091713 Arabidops
43	664.6	18.3	33563	8	AT1086	AT1086 Arabidops
44	648.2	17.9	1515	8	AF323040	AF323040 Zinnia el
45	648	17.9	1699	8	AY059858	AY059858 Arabidops

ALIGNMENTS

RESULT 1  
AF200533  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
zeae mays.  
zeae mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 3795)  
Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,  
Xoconostle-Cazares,B. and Delmer,D.P.

AF200533 3795 bp mRNA linear PLN 31-AUG-2000  
zeae mays cellulose synthase-9 (Cesa-9) mRNA, complete cds.



Db 1594 CGCGGGCCATGAAGAGAGAAATATGAAGAATTCAAAAATTAGGGTAAATGGCCCTTGTTGCT 1653  
QY 1476 AAGCATTTGAAGTCCCGAGGAGGATGGATCATCAAGATGGCACACCATGGCCAGGA 1535  
Db 1654 AAGGCACAAAATGCTCCTGAGGAAGGATGGATCATCAAGATGGCACACCATGGCCAGGA 1713  
QY 1536 AACAATACAGGAGCATCCTCGAATGATTCAGGTTTTCTTGGTGCACAGTGGTGGCCCTT 1595  
Db 1714 AACAATACAGGAGCATCCTCGAATGATTCAGGTTTTCTTGGTGCACAGTGGTGGCTTT 1773  
QY 1596 GATACCTAGGGTAAATGAGCTCCCGCTTTAGTTATATGTCGTCGTCGTCGTCGTCGTCG 1655  
Db 1774 GATACCTAGGGTAAATGAGCTACCCGCTTGGTCTATGTTCTCGTGAAGAACGTCCTGGA 1833  
QY 1656 TTCAGCACACAAAGGCTGGTGCATGAATGCCCTTGTTCGTCCTCAGCTGTCCTT 1715  
Db 1834 TTCAGCATCACAAAGGCTGGTGCATGAATGCTTCTGTCGCGCTCAGCTGTCCTT 1893  
QY 1716 ACTAATGGCAATACATGTTGAATCTTGATTTGATCTACTATACATCAACACAGCAAGGCT 1775  
Db 1894 ACCAATGGCAATACATGTTGAATCTTGATTTGATCTACTATACATCAACACAGTAAGGCT 1953  
QY 1776 GTCCGGAACCTATGCTTCCTAATGGATCCAAACCTAGGTCGCGCAAGTCTGTTATGTC 1835  
Db 1954 CTCAGGAAGCTATGCTCTCTTATGGATCCCTAACTAGGAAGAGTGTCTGCTATGTT 2013  
QY 1836 CAGTTCACCAAAAGGTTGATGGGATTTAGGAATGATCGATATCAACACAGCAACACT 1895  
Db 2014 CAGTTCACCAAGGTTGATGGGATTTAGGAATGATCGATATCCACAGGAACACC 2073  
QY 1896 GTCTTTTTCATTAACCTGAGGGCCCTTGACGGCATTCAGGACACAGTTCATGTTGGA 1955  
Db 2074 GTGTTTTTCATTAACCTGAGAGGCTTGATGGCATCCAGGACCACTTATGTTGGC 2133  
QY 1956 ACTGGTTGTTTTCACAGAACAGCTATCTATGTTATGAGCCCCCAATTAAGCGGAAG 2015  
Db 2134 ACTGGTGTGTTTTCACAGAACAGCTCTATATGTTATGAGCCCCCAATTAAGCAAAAG 2193  
QY 2016 AAGCCAGGTTCTTGGCATCAGTATGTTGGGCAAGAACAGCAAGCAAGTCAAGAAA 2075  
Db 2194 AAGGGTGTCTTGTCATCAGTATGTTGGGCAAGAACAGCAAGTCAAGAAA 2252  
QY 2076 AGGAGCTCAGATAAGAAAAGTGCACAAAGCATGTGGACAGTTCCTGTCAGTATTCAT 2135  
Db 2253 --GGGCTCACAGAAAAGTGCACAAAGCATGTGGACAGTTCCTGTCAGTATTCAT 2310  
QY 2136 CTCGAACATAGAGAGGTTGTAAGGTCGTTGGTGTGATGATGAGAAATCAGTTCTC 2195  
Db 2311 CTTGAAGATATAGAGAGGAGTTGAAGGCGCTGGATTTGATGAGAAAATCAGTTCTT 2370  
QY 2196 ATGTCCTCAATGAGCTTAGAGAGAGATTTGGCCAGTCAGCAGCATTTGTCCTCCACT 2255  
Db 2371 ATGTCCTCAATGAGCTTAGAGAGAGATTTGGCCAACTGTCAGCTTTTGTGCGTCCACT 2430  
QY 2256 CTGATGGAATATGGTGTGTTCTCAGTCTCCACTCCAGAACTCTCTTTTGAAGAAGCT 2315  
Db 2431 CTGATGGAATATGGTGTGTTCTCAGTCTGCGACTCCAGAACTCTCTTGAAGAAGCT 2490  
QY 2316 ATCCATGCTATAAGTTGTGGCTATGAGGACAACTGTAATGGGAACTGAGATTTGGTGG 2375  
Db 2491 ATCCATGCTATAAGTTGTGGCTACGAGGACAAAGATTTGAATGGGAACTGAGATTTGGTGG 2550  
QY 2376 ATCTATGGATCTGCACAGAGATATCTTACTGGATTCAGATGCAACGACAGAGCTGG 2435  
Db 2551 ATCTATGGTCTGTGACGGAAGATATCTCAGTGGGTTCAAGATGCAACGACAGAGCTGG 2610  
QY 2436 CGTTCAGTCTATTGCTATGCCAAAGCCGACGCTTTTCAAGGAGTCTCCGCCCATCAATCTT 2495  
Db 2611 CGGTGATCTACTGCTATGCTTAGCGCGCCGCTTCAAGGGATCGCTCCCATCAATCTC 2670  
QY 2496 TCAGATGCTGAACCAAGTGCTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGC 2555  
Db 2671 TCAGACCGTCTGAACCAAGTGCTGCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTTCAGC 2730

QY 2556 CGGCATTCGCCCTTATGGTATGGCTACGGAGGGCGCCTCAAGTTCCTCGAGAGATTCGCT 2615  
Db 2731 CGGCATTCGCCCTTATGGTATGGCTACGGAGGACGCTTGAAGTTCCTGGAGAGATTCGCC 2790  
QY 2616 TACATCAACACACACATTTTACCCTACTTCTCCGCTTCTAGTCTATGTTATATG 2675  
Db 2791 TACATCAACACACACATCTACCGCTCAGTCCCTCCGCTCCTCATTTACTGTATCCTG 2850  
QY 2676 CCTGCTATCTGTCGCTCAGTGGAAAGTTCATGATCCAGAGATAGCAACTTGGCCACT 2735  
Db 2851 CCTGCTATCTGCTCTCAGCGGGAAGTTCATCATCCAGAGATCAGCAACTTCGCTAGT 2910  
QY 2736 ATCTGGTTCATTCGCTCTTCTTCAATTTTGGCACTGGTATCCTTGAGATGAGGTGG 2795  
Db 2911 ATCTGGTTCATCTCTCTCTTCTCATCTCGATCTTGGCCAGGCTATCCTGGAATGAGGTGG 2970  
QY 2796 AGTGTGTTGGATGACAGTGGTGGAGGAATGAACAGTTCCTGGGTCAATTTGGAGGTATC 2855  
Db 2971 AGCGGCTGGGCATCGAGGAGTGGTGGAGGAACGAGCAGTTCCTGGGTCTATCGGAGCATC 3030  
QY 2856 TCTGCACATCTGTTTGGCGCTTTTCAGGCTCTTCTGAAGTTCCTTGGCGGTATCGACACC 2915  
Db 3031 TCGCGCCACCTCTTCCGCGCTTTCAGGCGCTTCTCAAGGCTCTTGGCGGATCGACACC 3090  
QY 2916 AACTTCACTGTCACTCAAGGCTAATCAGGAAGAGGCGACTTTTGTGCTGAGCTCTACATG 2975  
Db 3091 AACTTCACTGTCACTCCAGGCTCGGATGAAGAGGCGGACTTCGCGGAGCTGTACATG 3150  
QY 2976 TTCAAGTGGAGCAGCTTCTCATCCCTCCGACGACATTTTGTATCATTAACATGTTGGT 3035  
Db 3151 TTCAAGTGGAGCAGCTTCTGATCCGCGCCACCACCATCTGATCATCAACTGGTCGSC 3210  
QY 3036 GTGCTGTGGCAGCTCTCTAGCCATCAACAGTGGTTPACCAATCATGCGGCGCGCTCTTT 3095  
Db 3211 GTTGTGTGGCGCATCTCTAGCCATCAACAGCGGCTACCACTGCTGGGGTCCGCTCTTC 3270  
QY 3096 GGAAGCTCTCTTCTTGGCTCTTGGGTGATTTTCACTTATACCCATTCCTCAAGGCTCTT 3155  
Db 3271 GCAAGCTCTCTTCTGCGCTTCTGGGTGATGTTTCACTGTACCGGTTCTCTCAAGGCTCTC 3330  
QY 3156 ATGGCAGGCAAAACCGCACACGATTTGTCTGCTGGGTCTCTCTCTCTCTCTCTCTCT 3215  
Db 3331 ATGGGTGCGCAGAACCGCACCCGACCATCTGTTGTTGCTGGGCGATCCTCTGTCGCTG 3390  
QY 3216 ATCTTCTCTCTGCTGGGTGCTGTTGATCCATTCATACCGTCTCTGCTGGCGCCAAAT 3275  
Db 3391 ATCTTCTCTCTGCTGGGTGCTGATCGATCCGTTTCAACAAACCGGCTCAGTGGCCGGAT 3450  
QY 3276 ATCCAAACCTGTGGCATCACTGTAGGAAAGTGGGA-GTTTTGTAGACACAGAAA 3329  
Db 3451 ACTGAACGTGGGCATCACTGTAGGAGGTGGAAAGTTTTGTAGAAACAGAGA 3505

RESULT 2  
AX338680  
LOCUS AX338680 3799 bp DNA linear PAT 09-JAN-2002  
DEFINITION Sequence 5 from Patent WO0179516.  
ACCESSION AX338680  
VERSION AX338680.1 GI:18129042  
KEYWORDS  
SOURCE Zea mays.  
ORGANISM Zea mays.  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
1  
REFERENCE  
AUTHORS Dhugga, K. S. and Helentjaris, T. G.  
TITLE Maize cellulose synthases and uses thereof  
JOURNAL Patent: WO 0179516-A 5 25-OCT-2001;  
PIONEER HI-BRED INTERNATIONAL, INC. (US)  
FEATURES  
Location/Qualifiers  
source 1. .3799

GAATGTTGCC	TGGAAAGAGAGATTGATGGCTGG	695
GAATGTTGCC	TGGAAAGAGAGGGTTGATGGCTGG	873
TC	CCATGACTAATGGGACAAAGCATTTGCTCCCTCT	755
TC	CCATGACGAATGGCACAGCATTTGCTCCCTCT	93
TGCATCTACT	CAATACAACTGGGAAGACGCTTTA	815
TGCATCAACTGAT	TACACATGGGAAGATGCCTTA	993
ATCTAGAAAAGT	CCCAATTGCTTCCTCCAAAATA	875
ATCTAGAAAAGT	CCACTTCCTCCAGGATA	1053
CGGTTGGTTGTTCT	TAGCATCTTCCTGCACATAC	935
ACGATTGATTGTTCT	TAAGCATCTTCTTGCACATAC	1113
ATACCCACTGTGGCT	TTTATCTGTTATATGTGAG	995
ATACCCACTGTGGCT	TCTATCTGTTATATGTGAG	1173
GGATCAGTTCCCGAAGTGGTTT	CCAATCAACCGG	1055
GGATCAGTTTCCAAAGTGGTTT	CCAATCAACCGG	1233

[illegible]





/product="cellulose synthase-4"  
/protein\_id="AAP89964.1"  
/db\_xref="GI:9622880"

/translation="MEGDADGVKSRGGQVCQICGQGVGTAEQDVFAACDVCGP  
VCPCPEYERKQDTPCQCKTKYKHKGPSAIRGEGDDTDADDFNYLASGNEQK  
QKIDMRWRNVNGSGDVGPRKYDSGEIGLTKYDSGEIPRGYIPSTVNSQISGEIP  
GASPDHMKSPGNTGKRAPFYVNHSPNSPREFSGSIGNVAKRERQVWKQDKGT  
IPWNTGTSAPSEGRGVGDIDASTDYNMEDALLNDETQPLSRKVPPLSSRINPYRMV  
IVLRILVLSIFLHYRITNPVRNAYPLMLLSVCEJFWFALSWILDOFPKWFPIRETYL  
DRLALYREGPESQLAADIVFVSDPMKEPLVANTVLSILAVDIPVDRVCYVS  
DGAAMLTDEALAESEFARKVPFVKNIIEPRAPETFSOKIDYLDKXVHPSPFXD  
RRAMKREYEFVRYVNGLVAKAQPEEIMQDTPWPNGNTRDHPQIMQVGLHSG  
GLMTEGNEPLRVLYNREKRPQFQHKKAGAMNALVRVSAVLTYNQVLMNDCDHYIN  
NSKALREACFMFLNPLGVCYVQPOFDFGIDRNDYANRNTFFDINLRGLDGIQ  
GVYVCTGCVENRTALYGEYPPIKQKGFSLSCGGRKAKSKKSDKKSOKHVD  
SSVPELNEDIEGVEGAGFDEKSLMSOMSLKREKFGQSAFVASTLMELMEYGGVQPSA  
TPSLKAEIHWISQEDKTEGWTEIGWYGVTEIDLTGFKMHARGWVYICMPKR  
PAPKGAENLDRNLQVLRWALGSEVILFSRHCPLWYGYGRKLFBRFYAINITY  
PLTSIELLYCILPAICLLTGKFIIPESNFASIFWISLFIATGILEMRWSGVI  
DEWRNEQFWIIGGISAHLFVAFQGLKVLAGIDTNFTVSKADEGDFAEIYMKW  
TTLIPPTTILINLVGVVAGISYAINSQYQSWGFLGKLPFAFWVIHLVPLKGLM  
GRONRPTIVVWAILLASIFSLWVRIDPFTTRVTPDPTQCGINC"

BASE COUNT 861 a 975 c 1006 g 902 t 1 others  
ORIGIN

Query Match 70.9%; Score 2572.2; DB 8; Length 3745;  
Best Local Similarity 86.4%; Pred. No. 0;  
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;

Qy	4	CGAGAACCCCGCTCGTCTGCTCGGTGCGGGTGGATCGCTCGCCGGCCCATGG	53
Db	264	CCAGGCCCGAGCTCCAGGCGACGCTCCCTCGACGTTCTCGGCGAGCTCGCTTGC	323
Qy	64	ACGGCGACGGCGCCCTGAAGTCGGGAGGACGCGGGCGGGGACGTGTGCGACATCT	123
Db	374	AGGGCGACGGCGCGTGAAGTCGGGAGGCGGGTGGCGACAGGTGTGCGACATCT	383
Qy	124	CGCCGCGACGGCGTGGGACACAGTTGAGCGGCGACGCTTCTACCGGCTCGACGTGCG	183
Db	384	CGCGCGACGGCGTGGGACACACGCGGAGGGGACGCTCTTCGGCGCTCGACGTGCG	443
Qy	184	GCTTCCGGCTCGCCGCTGTACGACGACGACGACGAGGAGGACCCAGGCGCTGCC	243
Db	444	GTTTCCGGTGTGCGCCGCTGTACGAGTACGAGGCGACGAGGAGGAGGCGGCTGCC	503
Qy	244	TCCAGTCAAGACCAAGTACAAGCGCCACAGAGGAGCGCCGACGATCCGCGGAGGAAG	303
Db	504	CCCAGTCAAGACCAAGTACAAGCGCCACAGGAGGAGCGCCGCGATCCCGTGGGAGGAAG	563
Qy	304	CGGACGACATGATGCGCGATGTAGTACCTTCAACTACCTGATCTGGCAGTGAGG	363
Db	564	GAGACGACATGATGCGCGATGTAGTACCTTCAACTACCTGATCTGGCAGTGAAG	617
Qy	364	ACCAGAGCAGAGGATGCTGACAGGATGCGCAGCTGGCGCATGAACCCGGGGCGAGTG	423
Db	618	ACCAGAGCAGAGATGCTGCGCAGAGATGCGCAGCTGGCGCATGAACCTGGGGCAGCG	677
Qy	424	GCAATGTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTTCCAAAGTATGACAGTG	483
Db	678	GGGATGTGGTGGCCCAAGTATGACAGTGGCGAGATCGGGCTTACCAAGTATGACAGTG	737
Qy	484	GAGAGATCCCTTAGGGATAGCTCCCTTCAGTCAACACGACGACGATGTGAGGAGAAATCC	543
Db	738	CGGAGATTCCTTCGGGGATACATCCCATCAGTCACTAACACGACGATCTCAGGAGAAATCC	797
Qy	544	CTGAGCTTCGCTGTATCATCATGATGTCCCTTACCGGGAAACATCAGCAGACGTGCTC	603
Db	798	CTGCTGTTCCTTCGCTGTATCATCATGATGTCCCTTACCGGGAAACATCAGCAGACGTGCTC	857
Qy	604	CGTTTCCCTATGTGAATCATTCACCAATCCGTCAAGGGAGTTCTCCGGCAGTATTTGGGA	663
Db	858	CATTTCCTATGTGAACCATTCGCCAAATCCCGTCAAGGGAGTTCTCTCTGTAGCATTTGGGA	917

Qy	664	ATGTTGCTGGAAAGAGAGAGAT	723
Db	918	ATGTTGCTGGAAAGAGAGAGAT	977
Qy	724	CCATGACTAATGGGACAAAGCA	783
Db	978	CCATGACCAATGGGACAAAGCA	1037
Qy	784	CATCTACTGAATACAACATGGA	843
Db	1038	CATCAACTGAATACAACATGGA	1097
Qy	844	CTAGAAAAGTCCCATTCGCTTC	903
Db	1098	CTAGAAAAGTCCCATTCGCTTC	1157
Qy	904	GTTTGGTGTCTTAAGCATCTT	963
Db	1158	GATTGATGTCTTAAGCATCTT	1217
Qy	964	ACCCACTGTGGCTTTTATCTGT	1023
Db	1218	ACCCATATGGCTTCTATCTGT	1277
Qy	1024	ATCAGTCCCGAAGTGGTTTCC	1083
Db	1278	ATCAGTCCCGTAACTGGTTTCC	1337
Qy	1084	GGTATGACCGGAGAGGTCACCC	1143
Db	1338	GGTATGACCGGAGAGGTCACCC	1397
Qy	1144	TCGACCCCTTTGAAGGAGCCACC	1203
Db	1398	TCGACCCCAATGAAGAGCCACC	1457
Qy	1204	TTGATATTCGCTGGGACAGGTT	1263
Db	1458	TGGATATTCCTCTGGTAAAGTT	1517
Qy	1264	CTTTTGGACGATTCGGCTGAGAC	1323
Db	1518	CATTGATGCACTAGCTGAGAC	1577
Qy	1324	AGTATGATTAATGAACCCAGAC	1383
Db	1578	AGTACAACATTTGAACCTAGAC	1637
Qy	1384	AAGACAAATCCAGCCTTCATT	1443
Db	1638	AGGACAAATGTCACCTTCATT	1697
Qy	1444	AATTTAAATCAGGATAAATGCA	1503
Db	1698	AATTCAAAGTTAGGTAATGGA	1757
Qy	1504	GGATATGCAAGATGGCACACC	1563
Db	1758	GGATATGCAAGATGGCACACC	1817
Qy	1564	TTCAAGTGTTCCTTGGTTCACAG	1623
Db	1818	TTCAAGTGTTCCTTGGTTCACAG	1877
Qy	1624	TAGTTTATGTCTCTGTGAAAG	1683
Db	1878	TGGTCTATGTCTCTGTGAAAG	1937
Qy	1684	TGAATGCCCTGTCTGTCTCTC	1743
Db	1938	TGAATGCCCTGTCTGTCTCTC	1997
Qy	1744	ATTGTGATCTACTACATCAACA	1803

Db 1998 ATTGTGATCACTACATTAACAAAGTAAGGCTCTCAGGGAAGATATGTGCTTCCTTATGG 2057  
Qy 1804 ATCCAAACCTAGGTCCGCAAGTCTGTATTATGTGCAGTTCACCAAGGTTTGATGGGATGG 1863  
Db 2058 ACCCTAACCTAGGAAGAGTGTCTGCTACGTCCAGTTCCTCCAGAGATTCCGATGGCATGG 2117  
Qy 1864 ATAGGAATGATCGATATGCAAAACAGGACACTGTCTTTTGTATATTAATCTTGAGGGGCC 1923  
Db 2118 ACAGGAATGATCGATATGCCAACAGGAACACCGTGTCTTTCGATATTAATCTTGAGAGGTC 2177  
Qy 1924 TTGACGGCATTCAGGACCAAGTATTGTGGAACTGGTGTCTTTTCAACAGAACAGCTA 1983  
Db 2178 TTGATGGCATCCAAAGGACCAAGTATTGTGCGAACTGGCTGTGTTTCAACCCGAACAGCTC 2237  
Qy 1984 TCTATGTTTATGAGCCCCCAATTAAGCGGAAGACCGAGTTCCTTGGCATCACTATGG 2043  
Db 2238 TATATGGTTATGAGCCCCCAATTAAGCAGAAGAGGTGTTCTTGTGCATCACTATGG 2297  
Qy 2044 GGGCAAGAAAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAGAAAAGTCAACA 2103  
Db 2298 GCGGTAGGAAGAGGCAAGCAAAATCAAAAGAA---GGGCTCGGCAAGAAAGTCCGAGA 2354  
Qy 2104 AGCATGTGGACAGTCTGTTCACAGTATTCAATCTCGAAGACATAGAGGAGGTGTTGAAG 2163  
Db 2355 AGCATGTGGACAGTCTGTGCCAGTATTCAACCTTGAAGATATAGAGGAGGAGTTGAAG 2414  
Qy 2164 GTGCTGGGTTTGATGATGAGAAATCAGTCTCATGTCTCAATAGCTTAGAAGAGAGAT 2223  
Db 2415 GCGCTGGATTTGACGAGAGAAATCACTTCTTATGCTCAAAATGAGCCTGGAGAGAGAT 2474  
Qy 2224 TTGGCCAGTCAGACAGATTTGTGGCTCCACTCTGTAGGAATATGTGGTGTTCCTCAGT 2283  
Db 2475 TTGGCCAGTCGACAGCTTTGTGGCTCCACTCTGTAGGAGTATGTGGTGTTCCTCAGT 2534  
Qy 2284 CCTCCATCCAGAAATCTCTTTTGAAGAAGCTATCCATGTATCAATGTGTGGCTATGAGG 2343  
Db 2535 CCGCAACTCCGGAGTCTCTTCTGAAGAAGCTATCCATGTATCAAGTGTGGCTATGAGG 2594  
Qy 2344 ACAAGTCTGATGGGAACTGAGATTTGGTGGATCTATGATCTGTACAGAGATATTC 2403  
Db 2595 ACAAGTCTGATGGGAACTGAGATCGGGTGGATCTACGGTCTGTGACAGAGAGATTC 2654  
Qy 2404 TTACTGGATCAAGATCAGCAGAGGCTGGGTTTCAGTCTATTGATGCCCAAGCGCC 2463  
Db 2655 TCACCGGATTCAGATCACGCGGAGGCTGGCGGTGATCTACTGATGCCCAAGCGGC 2714  
Qy 2464 CAGCTTTCAAGGGATCTGCCCCATCAATCTTTCAGATCTGTGAACCAAGTCTCGGT 2523  
Db 2715 CAGCTTTCAAGGGTCTGCCCCATCAATCTTTCGGACCGCTGTGAACCAAGTCTCGGT 2774  
Qy 2524 GGGCTCTCGGTTCTGTTGAATCTTTTACGGCGCAATGCCCTTATGATGGCTACG 2583  
Db 2775 GGGCTCTGGTCTCGTGAGATCTCTTACGGCGCACTGCCCTGTGTGATGGCTACG 2834  
Qy 2584 GAGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCACTTACCCACTAA 2643  
Db 2835 GAGGGCGCTCAAGTCTCTGAGAGATTCGCTTACATCAACACCACTTACCCGCTCA 2894  
Qy 2644 CCTCTCTCCGCTCTAGTCTATTGTATATGCTGTATCTGTCTGCTCAGTGGAAAT 2703  
Db 2895 CGTCCATCCCGCTCTCATCTACTGCACTCTGCCCGGCATCTGTCTGCTCACCGGAAGT 2954  
Qy 2704 TCATCATGCCAGAGATTAGCAACTTGGCAGTATCTGTTTCATTGGCTCTCTCTTCA 2763  
Db 2955 TCATCATGCCAGATCAGCAACTTCCAGCATCTGTTTCATCTCCCTCTTTCATCTCGA 3014  
Qy 2764 TTTTCGGCACTGGTATCTCTGAGATGAGTGGAGTGGTGTGGCATTTGACGAGTGGTGA 2823  
Db 3015 TCTTCGGCACGGGATCTCTGAGATGAGTGGAGCGGGGTGGGATCGAGAGTGGTGA 3074  
Qy 2824 GGAATGAACAGTCTCTGGGTCTATTGGAGGTATCTCTGCACATCTTGTTCGGCTTTTCAG 2883

Db 3075 GGAACGACGAGTTCGTGGTGATCGGGGCAATCTCCGGCGACCTCTTCGCCCGTGTTCAGG 3134  
Qy 2884 GTCTTCTGAAGTGTGTTGCCGTATCGACACCAACTTCACGTGTACCTCAAGGCTAATG 2943  
Db 3135 GCTCTCTCAAGTGTGTCGGCGCATCGACACCAACTTCACCGTCACTCCCAAGGCTCGG 3194  
Qy 2944 AGGAAGAAGCGGACTTTCCTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCCTC 3003  
Db 3195 ACGAGGACGGGAGCTTCGCGGAGCTGTACATGTTCAAGTGGACGACGCTTCTGATCCCGC 3254  
Qy 3004 CGACACCATTTTGTATCATTAACATGTTGGTGTGCTGGCGACCTTCCTACGCCATCA 3063  
Db 3255 CCACACCATCTGTATCATCAACCTGGTTCGGGCTGCTGCCGCGCATCTCTTACGCCATCA 3314  
Qy 3064 ACAGTGGTTACCAATCATGAGGGGCGCTCTTTTGGGAAGCTCTTCTTTGCCCTTCCTGGGTGA 3123  
Db 3315 ACAGCGGATACGAGTGTGGGCGCGCTCTTCGGCAAGCTCTTCTGCGCTTCCTGGGTCA 3374  
Qy 3124 TTGTTCACTTATACCATTCCTCAAGGCTTATGGCGAGCAAAACCGCACCGACGA 3183  
Db 3375 TCGTCCACTGTACCGTTCCTCAAGGGCTTCATGGGCGAGCAGAACCGCACCGACCA 3434  
Qy 3184 TTGTCATGCTGGGCTGTCTCTCGCTTCTATCTTCTCCTGTGTGGGTTCTGTGTTG 3243  
Db 3435 TCGTCTGCTGGGCGCATCTCTGCTGGCGTCCATCTTCTCTGCTGTGGGTTTCGCATCG 3494  
Qy 3244 ATCCATTCACATACCGCTCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACTGCTAGG 3303  
Db 3495 ACCCTTTCACACCGCGCTCACTGGCCGGATACCCAGACGCTGGCATCAACTGCTAGG 3554  
Qy 3304 AAAGTGGGAGTTT 3316  
Db 3555 GAAGTGGAAAGTT 3567

RESULT 4  
AF200529 3676 bp mRNA linear PLN 31-AUG-2000  
LOCUS Zea mays cellulose synthase-5 (Cesa-5) mRNA, complete cds.  
DEFINITION  
ACCESSION AF200529  
VERSION AF200529.1 GI:9622881  
KEYWORDS  
SOURCE Zea mays.  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 3676)  
AUTHORS Holland,N., Holland,D., Helentjaris,T., Dhugga,K.S.,  
Xoonostle-Cazares,B. and Delmer,D.P.  
Xoonostle-Cazares,B. and Delmer,D.P.  
TITLE A comparative analysis of the plant cellulose synthase (Cesa) gene  
family  
JOURNAL Plant Physiol. 123 (4), 1313-1324 (2000)  
MEDLINE 20398328  
PUBMED 10938350  
REFERENCE 2 (bases 1 to 3676)  
AUTHORS Dhugga,K.S. and Helentjaris,T.G.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1999) Agronomic Traits, Pioneer Hi-Bred  
International, Inc., 7250 NW 62nd Avenue, Johnston, IA 50131, USA  
FEATURES  
source Location/Qualifiers  
1..3676  
/organism="Zea mays"  
/db\_xref="taxon:4577"  
/note="identified and isolated from the maize genome  
database cDNA collection of Pioneer Hi-Bred International,  
Inc"  
1..3676  
/gene="Cesa-5"  
254..3484  
/gene="Cesa-5"  
/note="CESA-5"  
/codon\_start=1

/product="cellulose synthase-5"  
/protein\_id="AAF89965.1"  
/db\_xref="GI:9622882"  
/translation="MDGDATNSGKHVAGOVQICGDDGVGTAAADGDLFTACDVCFFPV  
CRPCYEYERKDTQPCOCKYIKRHKSPVHGEENEDVDADDYDNYNQSNGDDQ  
KOKTAERLWRTNSRGSDIGLKYDSEIGEDIPRGYIPSDYTHSISGEIP  
GASPDHMSPVNIGRRHQFPYVNHSPREFSGSLGNVAKRVEDGWMKDKGAI  
PMTNITSAPSEGRVADIDASTDYNMEDALLNDTROPLSRKVPSPSRINPWRVI  
VLRLAVLCIFLRYRTHPVNNAYPLWLKVEICEIFWALSILWLDQPKVSPINREYLD  
RLALYREGPEQSQAPVDIFVSTVDPMKPEPLVANTVLSILAVDYPVKVSCSYSD  
DGAAMLTDAUSESEFAKRWPFCKKNIEPAPWYFAQKIDILDKOKVQVFSFKER  
RAMREYEFKRVINGLVAQKQPEEGIMQDGPWPGNNTDRHPGMQVFLHSGG  
LDVEKNEPLRVVYREKRPFGQHHKAGAMNALVRSAVLNNGYMLNLDCHYINN  
SKALREACFLMDPLNRYVQVQPPQDFDIDRNDYANRNTVFEDINLRGLDGIQ  
PVYVCTGVENRTALYVEPPVKKKPGFSLCGRRKTSKSKSKSHRADS  
SYPVNLNDEIEGSGQFDEKSLMSQMSLEKRFQSSVFASTLBYGQVPOSAT  
PESLKEAHINLSRDLNOVLWALGSIEILFHSRCPHWYGGRLKLEFPAYINCEPKRP  
AFKGSNPLNDRVCLLTGKFIIPKISNLESWFISLFIATGILMRWVSGVIGD  
EMRNEQFWIGGISAHLFAVFOGLKVLGIDTFTVTSKATDEEGFAELMEKWT  
TLIPPTLILNLNVVAGISYAINSGVQSWGLFGLFFAFWLVHLYPLKGLMG  
KONRPTIIVVWAILLASIFSLMWVRIDPFTTRVTPDIAKGCING"  
BASE COUNT 907 a 826 c 951 g 991 t 1 others  
ORIGIN

Query Match 64.1%; Score 2325.2; DB 8; Length 3676;  
Best Local Similarity 82.8%; Pred. No. 0;  
Matches 2706; Conservative 0; Mismatches 549; Indels 12; Gaps 4;

QY	38	GGTTGGATGCTGTGCGCGCCCATGACCGGACGCGGACGCCCTGAAAGTCCGGGAGGCA	97
DB	232	GGAGGGCTCGGATTTGGTGCCCATGACGCGGCGC---GACGCCACGAATTCGGGGAGCA	288
QY	98	CGGGCGCGGAGGTGTCCAGATCTGCGCCGACGCGCTGGGACACACCTGTGGACGCGGA	157
DB	289	TGTGGCGCGGAGGTGTCCAGATCTGCGCGGACGCGCTGGGACACACCTGTGGACGCGGA	348
QY	158	CGTCTTACCGCTGCGGAGCTGTCCCGCTTCCCGCTGTCGCGCCCTGTACACGACGCA	217
DB	349	CCTCTTACCGCTGCGGAGCTGTCCCGCTTCCCGCTGTCGCGCCCTGTACACGACGCA	408
QY	218	GCACAGGAGGACCCAGCGCTGCTCAGTGCACAGCAAGCAATCAAGCGCCACAGAGG	277
DB	409	GCACAGGAGGACCCAGCGCTGCGCGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	468
QY	278	GAGCCAGCGATCCGCGGGAGGAGGCGGACGACACTGATGCGCGATGATGATGATGATG	337
DB	469	GAGCCAGCGATCCGCGGGAGGAGGCGGACGACACTGATGCGCGATGATGATGATGATG	528
QY	338	CAACTACCTGCTGCTGCGACTGAGGACCAAGAGAGAGATGCTGACAGGATGCGCAG	397
DB	529	CAACTACCAAGCATCTGGCAACAGGATCAGAAAGCAAGATGCTGAGAGATGCTCAC	588
QY	398	CTGGCGCATGAACACCGGGGAGTGGCATGTTGGCCACCCCAAGTATGACAGTGGCGA	457
DB	589	TTGGCGGACAACTCAGTGGCAGTG---AATTTGGCTGGCTTAAATATGACAGCGGTGA	645
QY	458	GATCGGCTCTCCAACTATGACAGTGGAGAGATCCCTAGGGGATACGCTCCCTTCAGTCA	517
DB	646	AATTGGGCATGGAGATGACAGTGGTGAATCCCTCGTGATATATCCCGTCACTAAC	705
QY	518	CAACAGCCAGATGTACAGGAGAAATCCCTGGAGCTTCGCTGATCATCATCATATGTC	577
DB	706	TCATAGCCAGATCTCAGGAGAGATCTCTGGAGCTTCCCTGATCATCATCATATGTC	765
QY	578	TACGGGGAACATCAGCAGAGTCTCGTTCCTTCTGTAATCAATTCACCAATCCGTC	637
DB	766	TGGGAACATTCGGCAGCGTGGACATCAATTTCTTATGTAATCAATTCCTCAACCCATC	825
QY	638	AAGGGAGTTCTCCGGAGTATTCGGGAATGTTGCTGCGTGAAGAGAGAGTGTGATGCTG	697
DB	826	GAGGAGTTCTCCGGTGGCTTGGCAATGTTGCATGGAAGAGAGGTTGATGATGGA	885

QY	698	AATGAAGCAGGACAAGGGTGGC	757
DB	886	AATGA---AGGATAAAGTGC	942
QY	758	AGTCCGGCAGCTACTGACATC	817
DB	943	AGGCGTGGAGTTGCTGATAT	1002
QY	818	GAATGATGAACCTCGCAGCT	877
DB	1003	GAATGATGAACCTCGCAGCT	1062
QY	878	TCCTACAGAAATGGTCATGTT	937
DB	1063	TCCTACAGAAATGGTCATGTT	1122
QY	938	TCACAAATCCCTGCGGTAA	997
DB	1123	TATCACACATCCCTGTAACAA	1182
QY	998	TTGGTTTGTCTTATCCCTGGAT	1057
DB	1183	CTGGTTTGTCTTATCCCTGGAT	1242
QY	1058	GACCTACCTTGATAGACTGGC	1117
DB	1243	AACATACCTTGATAGACTGGC	1302
QY	1118	TGCTGTTGACATATTTGTGAG	1177
DB	1303	TCCTGTTGATATTTTGTGAG	1362
QY	1178	CAACACTGTCTTATCCATCTT	1237
DB	1363	AAATACCTGTCTTATCCATCTT	1422
QY	1238	ATCTGATGACGAGCTTCAATG	1297
DB	1423	TTCCGATGATGAGCTGCTATG	1482
QY	1298	TAGGAATGGTACCATTGCTG	1357
DB	1483	TAGGAATGGTACCATTGCTG	1542
QY	1358	CTTTTGCAGAAAATGATTAG	1417
DB	1543	CTTTTGCAGAAAATGATTAG	1602
QY	1418	CCGGGCCATGAAGAGAAATG	1477
DB	1603	CCGGGCCATGAAGAGAAATG	1662
QY	1478	GGCATGAAAGTCCCGAGGA	1537
DB	1663	GGCATGAAAGTCCCGAGGA	1722
QY	1538	CAATACCGGATCATCCTGG	1597
DB	1723	CAATACCGGATCATCCTGG	1782
QY	1598	TACTGAGGTAATGAGCTCCG	1657
DB	1783	CGTTGAAGCAATGAATCTCC	1842
QY	1658	CCAGCACCAAGAGCTGG	1717
DB	1843	CCAGCACCAAGAGCTGG	1902
QY	1718	TAATGGAATATCATGTTGAA	1777
DB	1903	TAATGGAATATCATGTTGAA	1962
QY	1778	CCGAGAAGCTATGTGCTCCT	1837

Db	1963	TCGAGAAAGCTATGTGCTTCCTTATATGGAGCCCAAACCTAGGAAGGAATGTCTGTTATGTGCCA	2022
Qy	1838	GTTCACCAAAAGTTTGATGGGATTTAGAGAAATCATCGATATGCAACAGGACACTGT	1897
Db	2023	ATTTCCTCAGAGGTTTGATGGTATTTAGTAGGAATGACCGATATGCAACAGGAACTGT	2082
Qy	1898	CTTTTGTGATATTAACCTTGAGGGGCTTGAGGGCATTCAGAGCATTCAGAGCAGTTATGTGGGAAC	1957
Db	2083	GTTTTTCGATATTAACCTTGAGAGGCTTGACGGCATTCAGGGCCAGTTATGTGGGAAC	2142
Qy	1958	TGTTTGTGTTTTCACAGAACAGCTATCTATGTGTTATGAGCCCCCAATTTAAGGGGAAGAA	2017
Db	2143	TGTTTGTGTTTAAACAGAACGCCCTTATATGGTTATGAGCCTCCAGCTCAAGAAAAAAA	2202
Qy	2018	GCCAGGTTTCTTGGCATCACTATGTGGGGCAAGGAAGCAAGCAAGTCAAGAAAAAG	2077
Db	2203	GCCAGGCTTCTTCTCTTCGCTTTGTGGGGGAAGGAAGACGTCAAAATCTAAGAA---	2259
Qy	2078	GAGCTCAGATAAGAAAAAGTCGAACAAGCATGTGGACAGTTCTGTTCGAGTATTCAATCT	2137
Db	2260	GAGCTCGGAAAAGAAGACTCACATAGACACGACACAGTTCTGTACCAGTATTTAATCT	2319
Qy	2138	CGAAGACATAGAGGAGGGTTTGAAGGTGCTGGGTTTCATCATGAGAAAAATCAGTTCAT	2197
Db	2320	CGAAGATATAGAGGAAGGATTTGAAGGTTCTCAGTTTTCATGATGAGAAATCGCTGATTAT	2379
Qy	2198	GTCTCAATGAGCTTAGAGAAGAGATTTGGCCAGTCAGCAGCATTTGTTGGCTCCACTCT	2257
Db	2380	GTCTCAATGAGCTTGAGAGAAGATTTGGCCAGTCAGTGTTTTGTAGCCTCTACTCT	2439
Qy	2258	GATGGAATATGSGTGTTCCTCAGTCCTCCACTCCAGAAATCTCTTTTGAAGAAGCTAT	2317
Db	2440	GATGGAATATGSGTGTTCCTCAAAATCTCGAACTCCAGAGTCTCTTCTGAAAGAAGCTAT	2499
Qy	2318	CCATGTCATAAGTTTGGCTATGAGGACAAGTCTCAATGGGGAACATGAGATTTGGTTGGAT	2377
Db	2500	TCATGTCAATCAGCTGTGGCTATGAGGACAAAACACTGACTGGGSAACATGAGATTTGGGTGGAT	2559
Qy	2378	CTATGGATCTGTCCACAAGAAGATTTCTTACTGGATTCAGATGACGCAAGAGGCTGGCG	2437
Db	2560	CTATGGTTCTGTACAGAAGACATCTCACCGGATTCAGATGCAATGCTCAGGCTGGCG	2619
Qy	2438	TTCAGTCTATTGTCATCCCAAGCGCCAGCTTTCAGGGATCTGCCCCCATCAATCTTTC	2497
Db	2620	ATCAATCTACTGCATGCCTAAGCGACAGCTTTCAGGGATCTGCTCTCATCAACCTTTC	2679
Qy	2498	AGATCGTCTGAACCAAGTGTGGGTGGGCTCTCGGTCTGTGAAATCTTTTTCAGCGC	2557
Db	2680	GGATCGTTTGAATCAAGTGTCTGGGTGGGCTCTTGGTTCCATTTGAAATCTTTTTCAGCAG	2739
Qy	2558	GCATTTGCCCTTATGTGTATGGTACGAGGGCGCCTCAAGTTTCTGTGAGAGATTCGCTTA	2617
Db	2740	GCATTTGCCCATATGGTATGGCTATGGAGGGCGGCTTAAATTCCTGGAGAGATTTGCTTA	2799
Qy	2618	CATCAACACCAACATTATCCCACTAACCTCTCTCCGGTCTTAGTCTATTGTATATTGCC	2677
Db	2800	TATCAACACAAACATTTATCCACTCACATCAATCCCGCTCCTCCTGTACTGCATATTGCC	2859
Qy	2678	TGCTATCTGTCTGCTCACTGGAAAGTTCAATCATGCGCAGAGATTAGCAACTTTGGCCAGTAT	2737
Db	2860	AGCAGTTTGTCTTCTCACTGGGAAGTTCAATCATCCCAAGATTTAGTAACCTTAGAGAGTGT	2919
Qy	2738	CTGGTTCAATTCGCGCTCTTCTTCAATTTTCCCACTGGTATCCCTTGAGATCAGGTGGAG	2797
Db	2920	TTGGTTTATATCCTCTTTATCTCAATTTTCCCACTGGTATTCCTTTGAGATGAGGTGGAG	2979
Qy	2798	TGSGTTTGGCATTGACAGAGTGGTGGAGGAATGAACAGTTCTGGGTCAATTTGGAGGTATCTC	2857
Db	2980	TGGTGTGGCAATTCATGAATGGTGGAGGAACAGCAGTTCTGGGTCAATTTGGTGTATTTC	3039
Qy	2858	TGCACATCTGTTGCCGTCTTTCAGGGTCTTCTGAAGGTGCTTGCCTGGTATCGACACCAA	2917

Db	3040	TGCGCATTTATTTGCCCGTCTTCCAGGGGTCTCCTTGAAGGTGCTTGCTGTGATCGACACGAG	3099
QY	2918	CTTCACGTGTACCTCAAGGCTTAATGACGAAGAAGCGACTTGTGTGAGCTCTACATGTT	2977
Db	3100	CTTCACGTGTACCTCTAAGGCCACTGACGAAGAAGGTGATTTGCCGAGCTCTACATGTT	3159
QY	2978	CAATGGACGACGCTTCTCATCCCTCCGACGACCATTTTGTATCATTAACATGGTTGGTGT	3037
Db	3160	CAAGTGGACAACGGTCTCTGATGCCACCAACACCATATTTTGATCATCAACCTCGGCGGT	3219
QY	3038	CGTTGCTGGCACCTCCCTACGCCCATCAACAGTGGTTACCAAAATCATGGGGGCCGCTCTTTGG	3097
Db	3220	GGTGCCTGGCATTTCCCTACGCAATCAATAGCGGTTACCAAGTCATGGGGACCTCTTTTCGG	3279
QY	3098	GAAGCTCTCTTTTGGCTTCTGGGTGATTTGTTCACTATATACCATTCCTCAAGGGTCTTAT	3157
Db	3280	GAAGCTCTCTTTGGTCTGGGTGATTTGTTCCACCTGTACCCCTTCCTCAAGGGCTCAT	3339
QY	3158	GGGACGCAAAACCGCACACGAGATTTGTCATCGTCTGGGTGTCCTTCCTCGCTTCTAT	3217
Db	3340	GGGGAAGCAGAACCGCACGCGACCATTTGCTGCTGGGTATTCCTTCCTTCGCTCGAT	3399
QY	3218	CTTCTCCTTCGTGGGGTTCGTTGTATCCATTCATCCCGTCTCGCTGGCCCAAAAT	3277
Db	3400	CTTTTCCCTGATGGGTTCGTATCGATTCATTCACCCGCTGCTGGCCCTCGATAT	3459
QY	3278	CCAAACCTGTGGCATCAACTGCTAGGA	3304
Db	3460	CGCGAAATGTGGCATCAACTGCTAGGA	3486
RESULT 5			
AF150630			
LOCUS			
DEFINITION	AF150630	Gossypium hirsutum cellulose synthase catalytic subunit (celA3)	linear mRNA
ACCESSION	AF150630	AF200453	mRNA, complete cds.
VERSION	AF150630.2	GI:6446576	
KEYWORDS			
SOURCE			Gossypium hirsutum.
ORGANISM			Gossypium hirsutum
REFERENCE			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
AUTHORS			1 (bases 1 to 3723)
TITLE			Laosinchai, W., Cui, X. and Brown, R.M. Jr.
			A full length cDNA of cotton cellulose synthase has high homology with the Arabidopsis RSW1 gene and the cotton CelA gene (Accession No. AF200453) (PCR 00-002)
JOURNAL			Plant Physiol. 122 (1), 291 (2000)
REFERENCE			2 (bases 1 to 3723)
AUTHORS			Kimura, S., Laosinchai, W., Itoh, T., Cui, X. and Brown, R.M. Jr.
TITLE			Immunogold Labeling of Rosette Terminal Cellulose Synthesizing Complexes in a Vascular Plant (Vigna angularis)
JOURNAL			Unpublished
REFERENCE			3 (bases 1366 to 3470)
AUTHORS			Laosinchai, W. and Brown, R.M. Jr.
TITLE			Direct Submission
JOURNAL			Submitted (11-MAY-1999) Botany, University of Texas at Austin, W. 24th Street, Austin, TX 78713, USA
REFERENCE			4 (bases 1 to 3723)
AUTHORS			Laosinchai, W., Cui, X. and Brown, R.M. Jr.
TITLE			Direct Submission
JOURNAL			Submitted (17-NOV-1999) Molecular Genetics and Microbiology, University of Texas at Austin, Austin, TX 78712, USA
REMARK			Sequence update by submitter
COMMENT			On Nov 17, 1999 this sequence version replaced gi:5081778.
FEATURES			Location/Qualifiers
source			1. 3723
			/organism="Gossypium hirsutum"
			/cultivar="Texas marker-1"
			/db_xref="taxon:3635"
			/tissue_type="cotton fiber"

```

/dev_stage="24 days post anthesis"
1..3723
/gene="celA3"
CDS
269..3472
/gene="celA3"
/note="celA1 homolog"
/codon_start=1
/product="cellulose synthase catalytic subunit"
/protein_id="A039534.2"
/db_xref="GI:6446577"
/translation="MESEGDIGKPMKNGGQTCQICGDNVGNKNTDGPFIACNICAF
PVRCPCYEYRKDQNSQPOCKRYKWKSPAILGDRGTGGDADGASDFISENQE
QOKRLAGRWKNAKYGRGDEGVATYDKIEISHNIPLLTSGOEVSSELASAPERLS
MASPGVAGKSSIRVDPVREFSGSLGNVAKWERYDGMKOEKNTVPMSTCOATSE
RGLCDIDASTDVLDSQNDDEARQPLSRKVSYSVSSKINPYRMVILRLVILFLHY
RINYPNVALWHLISVICIWFPAISWILDQFPKWLVPVNRETYLDRLALRYDREGEPS
ELAAVDIFVSVDPLKEPLVLTANTVLSILAVDIPVKVSCYVSDSDGAAMLTEALSE
TSFARQVPCFKKYNIEPRAPEMYFAQKIDYLDKQVTSFVKDRAMKREFEYKVR
INGLRAKQVPEGMIMQDTPWPNNTDHPGMIQVFLGSGGDLAEGNEIPLRVY
VSRKRFQGHKKAGAMNALVRVSAVLNNGAFLLNDCDHYNNSKALREAMEFLMD
PNLKQVYVQFQRFQDIDRNDYANRNVFFDINLRGLDGIQGVYVGTGVENRT
ALYGEYPLAPKHKRTGILSSLOGSRKSSKSKGSDKKSGKHVDSTVPVFNLED
IERGEVAGDEDEKSLMSOMSELEKRFQSAVAVASTLMENGVPQSAPEPTEILLKEAI
HVSQGEDTDMGSEIGWYGSVDIEDILGFRMHWGRSIIYCKMRPAPKSPAPIN
LSRLNGLRWALGSVEILFSRHCPIWYGYGRKLWLEFAYVNTTYPYPAIPLMY
CTLPVACLNFKEIIPQISNLASIFLSIFATGILMKWNGVIGIDQWRNEQFW
VIGVSAHLFAVFQGLLKVLAGIDTFTVTSKADEGDFAEIYMKWTLLIPPTTL
LIINLVGAGISVIVNGSYPGMLPLGKLFKFAFWIILHYPLKGLMGRQNTPTIV
VWNSILLASIFSLWLWRIDPFTTRVGPDEQGCINC"
BASE COUNT 1029 a 700 c 880 g 1114 t
ORIGIN
Query Match 46.9%; Score 1699.8; DB 8; Length 3723;
Best Local Similarity 72.6%; Pred. No. 0;
Matches 2346; Conservative 0; Mismatches 827; Indels 60; Gaps 9;
QY 84 AAGTCGGGAGCAGCGGCGGACGTGCGCAGATCTCGCGGAGCGGCTGGGCACC 143
DB 296 AAGCCCATAGAAATCTTGGTGGCCAGACATGCCAAATCTGTGTGACAAATGTTGGCAAA 355
QY 144 ACGTTGGACGGCAGCTTCCACCGCCTGCGACGTCTGCCCTTCCCGGTCTGCCGCCCC 203
DB 356 AATACAGATGGTATCCATCTTCAATATATGTGCAATTTCTCTGTATGCGAGCCT 415
QY 204 TGCTACGAGCAGCGGACGAGGCGCACCGGCTGCTCCAGTGCAGTGCAGAACCAAGTAC 263
DB 416 TGTTATGAATATGAAGGAGGAGGGAATCAGTCTTGCCTCAATGCGAAGCAGCATAC 475
QY 264 AAGCGCCACAGAGGAGCCCGCATCCGCGGGAGGAGCGACGACACTGATGCCGAT 323
DB 476 AAGTGGCAAAAGGCAGTCTGCTATTTCTTGAGACAGGAAACGGGTGGTATGCAGAT 535
QY 324 GATGCTAGTACTCACTACCTGCGATCTCGGCTAGGACGAGCAGCAAGCAATGCT 383
DB 536 GATG---GCGGCAGTGATTTCAATTCACAGAAACCAAGAGCAGCAGCAAGCAAACTTGCA 592
QY 384 GACAGGATGCGCAGCTGCGCATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAG 443
DB 593 GAAAGATGCGAGGCTGGAATGCAAGTATGGTAGAGGGAGGATGTTGGTCTCCACT 652
QY 444 TATGACAGTGGGATCGGCTCTCCAAATGATGACAGTGGAGAGATCCCTAGGGGATAC 503
DB 653 TAT-----GATAAAGAAATTTCTATAACCATATTCCTCTGCTCACCAGTGGGCAG 703
QY 504 GTCCCTTCAGTCAACCAAGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCCCTGATCAT 563
DB 704 G-----AGGTTCTGTTGAGCTATCTGCTGCAATCATCCTACCTGAGCGG 742
QY 564 CACATGATGCTCCCTACCGGGAACATCAGCAGACGTGCTCGGTTTCCCTATGTGAATCAT 623
DB 743 CTTTCTATGGCATCTCCTGG-----AGTTGCTGGTGGGAATCGAGTATAGGGTGT 795
QY 624 TCACCAAAATCCGTCAAGGAGGATTCCTCGGCGAGTATGGGAATGTTGCTCGGAAAGAGAGA 683
```

```

Db 796 GGATCCAGTCAGGAGGATTTGG 796
QY 684 GTTGATGGCTGGAAATGAAG 796
Db 854 GTTGATGGCTGGAAATGAAG 796
QY 744 ATTGCTCCCTCTGAAGTCGG 796
Db 908 CAGCTACTTCTGAA---AGAG 796
QY 804 GAAGACGCTTTTACTGTAATGAT 796
Db 965 GATGACTCTCAACTGAATGAT 796
QY 864 TCCTCCAAATAAATCCCTAC 796
Db 1025 TCATCTAAAATAAATCCCTAT 796
QY 924 TTCTGCACTACCGTCTTCACA 796
Db 1085 TTCTTGCACTACCGAATAACA 796
QY 984 GTTATATGTGAGATTTGGTTG 796
Db 1145 GTGATATGTGAGATTTGGTTG 796
QY 1044 CCAATCAACCGGAGACCTAC 796
Db 1205 CTTGTTAATCGTGAGGATAC 796
QY 1104 CCGTCTCAGTTGGCTGCTGTTG 796
Db 1265 CCGTCCGAATTAGCTGCAGTTC 796
QY 1164 CCTATCTGCTACTGCCAACACTG 796
Db 1325 CCACCTTGTGACGCCACACTG 796
QY 1224 GTCTCTTGTGATGATCTGATG 796
Db 1385 GTCTCTTGTGATGTTCTGATG 796
QY 1284 ACTTCAGAGTTTGTAGGAAAT 796
Db 1445 ACATCAGAGTTTGCAGGAAAT 796
QY 1344 GCTCCCGAGTTTACTTTTGC 796
Db 1505 GCTCCAGAAATGGTATTTTGCA 796
QY 1404 TTTGTTAAAGACCCCGCGGCA 796
Db 1565 TTTGTTAAAGATCGTAGAGCT 796
QY 1464 GCTCTAGTTTCTTAAGGCATTG 796
Db 1625 GGGCTTCTCGCAAGGACACAG 796
QY 1524 CCATGGCCAGAAACAATACCA 796
Db 1685 CCATGGCTGGTATAACACCA 796
QY 1584 AGTGGTGGCTTGTATCTGAGG 796
Db 1745 AGTGGGGCCCTGATGCTGAGG 796
QY 1644 AAGGCTCCTGGGTTCCAGCACC 796
Db 1805 AAGGCTCCAGGCTTCCCAACACC 796
QY 1704 TCAGCTCTCTTACTAATGAC 796
```



Db 1865 TCAGCAGTCCTTACTAATAAGGAGCTTCTTGTTAAATCTTGATTGTCATCACTACATAAAT 1924  
QY 1764 AACAGCAAGCGTCGAGAACTATGCTTCTTAAGTGATCCAAACCTAGGTCCGCAA 1823  
Db 1925 AATAGCAAGCGTTAAGGGAAGCTATGTTTCTTGATGGATCCAAACCTTGGAAAGCAA 1984  
QY 1824 GTCGTGATGTGACGTTCGCCAAAGGTTTGATGGGATGTAGGAATGATCGATATGCA 1883  
Db 1985 GTTGTGTACGTTCAAGTTTCCCTCAAGGTTTGATGTGATGATGCGAATGATCGATATGCC 2044  
QY 1884 AACAGCAACACTGTCTTTTGTGATATTAACTTGAGGGGCTTGACGGCAATCAAGACCA 1943  
Db 2045 AACAGAAATACATGTTCTTCTTGATATAAATTTGAGGGTCTGGATGGCAATCAAGGGCG 2104  
QY 1944 GTTATGTGGGAACGTGGTGTCTTCAACAGAACAGCTATCTATGTTATGACGCCCCA 2003  
Db 2105 GTTATGTGGGACGGATGTTTCAATAGAACAGCAATGTATGGATATGAACCCCC 2164  
QY 2004 ATTAAGCGGAAG-----AAGCCAGGTTTCTTGGCATCACTATGTGGGG---CAAGAAG 2054  
Db 2165 CTTAAGCCCAAGCATAGAAAACAGGATACCTATCTCATTTGCGGAGGTTCTCGTAAG 2224  
QY 2055 AAGGCAAGCAAGTCAAAAGAGAGCTCAGATAAGAAAAGTCAAGCAAGCATGTGAC 2114  
Db 2225 AAGAGTTCAAAATCAAGTAAAAAGGATCAGACAAAGAGAAATCTGGCAAGCATGTTGAT 2284  
QY 2115 AGTTCGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTTGTAAGGTGCTGGGTTT 2174  
Db 2285 TCTACTGTACCAAGTATTCATCTAGAGGATATAGAAGAGGTTGTAAGGGGCTGGGTTT 2344  
QY 2175 GATGATGAGAAATCAGTCTCATGCTCAAAATGAGCTTAGAAGAGATTTGGCCAGTCA 2234  
Db 2345 GATGATGAAAGTCATCTAGTATGTCACAAATGAGCCTGGAGAAACGGTTGGCCAGTCT 2404  
QY 2235 GCAGCAATTTGTCCTCCACTCTGATGGAATAGTGGTGTTCCTCAGTCCCTCCACTCCA 2294  
Db 2405 GCTGTTTGTGTCCTACACTTATGGAATAAGTGGTGTCCCTCAGTCTCTACCCCC 2464  
QY 2295 GAATCTTTTGAAGAAGCTATCCATGTCATAGTTGGTGTGATGAGGACAAAGTCTGAA 2354  
Db 2465 GAAACACTCTTAAGGAAGCAATTCATGTTATCAGCTGTGGATACGAAGACAAACAGAC 2524  
QY 2355 TGGGAACAGATGATGTTGGATCTATGGATCTGTACAGAGATATTTCTTACTGGAATC 2414  
Db 2525 TGGGAAGTCAGATTTGGTGGATATGTTCTGTACAGAAGATATTTCTTACTGGATC 2584  
QY 2415 AAGATGCACGAAGAGCGTGGCGTTTCAGTCTATTCATGCCAAGCGCCAGCTTTCAAG 2474  
Db 2585 AAAATGCATGCCGTGGTGGCGGTCAATTTACTGCATGCCAAAACGCCAGCCTTTAA 2644  
QY 2475 GGATCGCCCCCATCAATCTTCAGATCGTGTGACCAAGTGTGCGGTGGGCTCTCGGT 2534  
Db 2645 GTTCTGCTCCATCAATCTTTTCAGATCGTGTGAACCAAGTGTGCGATGGGCTCTTGGT 2704  
QY 2535 TCTGTGAAATCTTTTACGCGGATTCGCCCTTATGATGCTACGGAGGGCGCTC 2594  
Db 2705 TCTGTGGAATCTTTCTAGTAGGCATTCGCCCTATCTGGTATGTTATGTTAGGTAGGCTC 2764  
QY 2595 AAGTTCCTGGAGAGATTCGCTTACATCAACACCACTATTAACCCACTAACCTCTCCCG 2654  
Db 2765 AAGTGGCTTGAAGATTCGCATATGTTAAACAACCATCTACCCAGTCACTGCTATTCT 2824  
QY 2655 CTTCTAGTCTATGATATGCTGCTATCTGTCTGCTCACTGGAAAGTTTCATATGCA 2714  
Db 2825 CTTCTCATGCTACTGTACACTGCCAGCAGTCTGTCTGCTCACTAACAAAGTTCATTATCCA 2884  
QY 2715 GAGATTAGCAATTCGCCAGTATCTGGTTTCATTCGCTCTTCTTCAATTTTCGCCACT 2774  
Db 2885 CAGATTAGTAACCTTTCAGTAGGCATTCGCTTATATCCCTCTTCTCTCCACTCTTTCAACT 2944  
QY 2775 GGTATCTTGGATGAGGTGGGTTGTTGGCATTGACGAGTGGTGGAGGAATGAACAG 2834  
Db 2945 GGTATTTTAAAAATGAAGTGAATGGTGTGGGAATTTGACCAATGGTGGAGAAATGAACAG 3004

QY 2835 TTCTGGGTTCATTGGAGGTATCTCTGCACATCTGTTTCCCGCTCTTTCAGGGCTCTTGAAG 2894  
Db 3005 TTTTGGGTATCGGTGGTGTGTCGGCACATCTTTTCTGCTGTTTCCAAAGGCTTCTCAA 3064  
QY 2895 GTGCTTCCGGTATCGACACCAACTTCACTGTCAACCTCAAAAGGCTAATGACGAAGAGGC 2954  
Db 3065 GTTCTTCTGSCATTGACACTAACTTCACTGCTCACTTCAAAAGCATCAGACGAAGATGG 3124  
QY 2955 GACTTTCGTGAGCTCTACATGTTCAAGTGGACGAGCTTCTCATCTCCCTCCGACGACCAT 3014  
Db 3125 GATTTTCGTGAACCTTACATGTTCAAAATGGACACCTTCTCATCTCCCAACCACTCTC 3184  
QY 3015 TTGATCATTAACATGTTGTTGCTGTTGGCAGCTTCTACGCCATCAACAGTGGTTAC 3074  
Db 3185 CTATCATAAACTTGGTGGGGTTGTCAGGAATATCGTATGTCATCAATAGTGGCTAC 3244  
QY 3075 CAATCATGGGGCGGCTCTTTTGGGAAGCTCTCTTTTCCCTTCTGGGTGATTTGTTCACTTA 3134  
Db 3245 CAATCATGGGTCCCTCTTTGGTAAGCTATCTTTTGGCTTCTGGGTGATCATCCATCTC 3304  
QY 3135 TACCATTTCTCAAGGCTTATGGGACGAGGAAACCGCACACCGAGAGATGTCATCGTC 3194  
Db 3305 TACCCCTTCTTGAAGGTTTGTATGGGACGCGCAAAACCGTACACCTACCATTTGTTGTA 3364  
QY 3195 TGGGCTGTCTCCCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3254  
Db 3365 TGGTCAATTCGCTTGGCTCCATTTCTTTGTTGGGTGGGATCGATCCCTTCCACA 3424  
QY 3255 ACCGCTCTCGGTGCCCAAAATATCCAAACCTGTGGCATCAACTGCTAGGAAAG 3307  
Db 3425 ACCAGATTACCGGCGGATGTAGACAGTGTGGAATAAAGTGTTCAGAGAG 3477

RESULT 6  
LOCUS AF027174 3682 bp mRNA linear PLN 07-FEB-1998  
AF027174 Arabidopsis thaliana cellulose synthase catalytic subunit (Ath-B)  
DEFINITION mRNA, complete cds.  
ACCESSION AF027174  
VERSION AF027174.1 GI:2827142  
KEYWORDS Arabidopsis thaliana.  
SOURCE Arabidopsis thaliana.  
ORGANISM Arabidopsis thaliana.  
REFERENCE 1 (bases 1 to 3682)  
AUTHORS Arioli, T., Peng, L., Betzner, A.S., Burn, J., Witte, W., Herth, W., Camilleri, C., Hofte, H., Plazinski, J., Birch, R., Cork, A., Glover, J., Redmond, J., and Williamson, R.E.  
TITLE Molecular analysis of cellulose biosynthesis in Arabidopsis  
JOURNAL Science 279 (5351), 717-720 (1998)  
MEDLINE 98111412  
PUBMED 9445479  
REFERENCE 2 (bases 1 to 3682)  
AUTHORS Arioli, T.  
TITLE Direct Submission  
JOURNAL Submitted (29-SEP-1997) Plant Science Centre, Australian National University, Acton, Canberra, ACT 200, Australia  
FEATURES  
Location/Qualifiers  
source 1..3682  
/organism="Arabidopsis thaliana"  
/cultivar="Columbia"  
/db\_xref="taxon:3702"  
gene 1..3682  
/gene="Ath-B"  
CDS 256..3453  
/gene="Ath-B"  
/note="RSW1-like"  
/codon\_start=1  
/product="cellulose synthase catalytic subunit"  
/protein\_id="AAC39336.1"

	Query Match	46.2%	Score 1675.2;	DB 8;	Length 3682;
	Best Local Similarity	71.4%;	Pred. No. 0;	Mismatches	868;
	Matches 2348;	Conservative	0;	Indels	72;
	Gaps	8;			
QY	30	TCGGTGGCGGTTGATCGCTCTGCCGGCCCATGACGGCAGCGGACGCCTGAAGTCC	89		
Db	229	TGGGTGTGGAAACAATAAGTGACAATCGAATCCGAAGGGAAATCCGGGAAACCGG	288		
QY	90	GGGAGGACGGGGCGGGGACGTGCGCAGATTCGCCGACGCGCCCTGGGCACCACTGTG	149		
Db	289	ATGAGAACAATTGTCGCAGACTTGGCAGATCTGTAGTACAATGTGGCAACACTGTT	348		
QY	150	GACGGCAGCTCTTCACCCTGCGAGCTCTGCCGCTTCCCGCTCTGCCGCCCTGCTAC	209		
Db	349	GATGGATCGTTTTGTGGCTTGTGATATTGTTCAITCCCAGTTTTCGSCCTTGCTAC	408		
QY	210	GAGCACGAGCGCAGGAGGACCCAGCCCTGCCCTCAGTGCAGAACCAAGTACAAGGC	269		
Db	409	GAGTATGAGAGGAAGATGGGAATCAATCTTGTCCTCAGTGCAAAACGACATACAAGG	468		
QY	270	CACAGAGGAGCCAGCAGATCCCGGGGAGGAAGCGCACGACACTGATGCCGATGATGGT	329		
Db	469	CTCAAAGTAGTCTCGTATTCCTGGTGTAAGACGAGGATGCGTTAGCTGATGAAGGT	528		
QY	330	A---GTGACTTCAACTACCCTGCATCTGGCACTGAGGACCAAGACAGAGATTGCTGAC	386		
Db	529	ACTGTTGAGTTCACCTACCCCT-----CAGAGGAGAAATTTCAAG	570		
QY	387	AGGATGGCAGCTGGCGCATGAAACCCGGGGCAGTGGCAATGTTGGCCACCCCAAGTAT	446		
Db	571	CGGATGCTTGTGGTCATCTTACTCGTGGGAAGGAGGAAATGGGGAACCCCACTAT	630		
QY	447	GACAGTGGCAGATCGGCCCTCTCCAAGTATGACAGTGGAGAGATCCCTAGGGGATACGT	506		
Db	631	GATAAAGG-----TCTCTCAAAATCATCTTCTCGTCTCACGAGGACAGACAAG--	679		
QY	507	CTTTCAGTCAACACACCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCCTGATCATCAC	566		
Db	680	-----ATACTTCAGSAGAGTTTCTGCTGCCCTCACCTGAAAGCGCTC	720		
QY	567	ATGATGTCCTCTACGGGGAACATCAGCAGACGCTGCTCGGTTTCCT--ATGTAATCAT	623		
Db	721	TCGTATCTTCTACTATCGCTGGGGAAAGCGCTTCCCTATTTCATCAGATGCAATCAA	780		
QY	624	TCACCAATCCGTCAGGGAGTTCTCCGGCAGTATTGGGAATGTTGCCTGGAAGAGAGA	683		
Db	781	TCACCAATAGAAGGATTGTGGATCCTGTGGACTCGGGAATGTAGCTTGGAGGAGAGA	840		
QY	684	GTTGATGGCTGGAAATGAAGCAGGACAGGGTCCGATCCCCATGACTAATGGGACAGC	743		

QY 1824 GTCTGTTATGTCAGTTCACCAAGGTTTGATGGATTGATAGGAATGATCGATATGCA 1883  
DB 1969 GTTTGTTATGTTTCAGTTCACCAAGGTTTGATGGATTGATAGGAATGATCGATATGCT 2028  
QY 1884 AACAGGAACACTGCTCTTTTGGATTAATTAACCTTGAGGGGCTTGACGGCATTTCAAGGACCA 1943  
DB 2029 AATCGTAATACCGTGTCTTTGATATTAACCTTGAGAGGTTTAGATCGGATTCAAGACCT 2088  
QY 1944 GTTTATGTGGAACTCGTTGTTGTTTCAACAGAACAGCTATCTATGTTTATGAGCCCCA 2003  
DB 2089 GTATATGTCGGAACCTGGATGCTGTTTCAACAGAACAGCATATACGCTTATGAACCTCA 2148  
QY 2004 ATTAAGS- ----CGRAGAGCCAGGTTTCTTGGCATCACTATGTTGGGGGCAAGRAG 2057  
DB 2149 ATAAAGTAAACACAGAAGCAAGCTTTTATCTTAAGCTCTGTGGTGGATCAAGAA 2208  
QY 2058 GCAAGCAAGTCAAGAAAGAGCTCAGATAAGAAAAAGTCGAACAAGCATGTGGACAGT 2117  
DB 2209 AAGAATCCAAAGCTAAGAAAGAGTCGGACAAAAAAGAAATCAGGACGATCTACTCA 2268  
QY 2118 TCTGTTCCAGTATTCATCTCGAAGACATAGAGGAGGTTTGAAGGTGCTGGGTTTGAT 2177  
DB 2269 ACTGTTCTGTATTCATCAACTCGATGATAGAGAGGAGTTGAAGGTGCTGGTTTGAT 2328  
QY 2178 GATGAGAAATCAGTCTCTATGCTCAAAATGAGCTTAGAGAAGAGATTGSCCAGTCAGCA 2237  
DB 2329 GATGAAGCGCTTTAATGTCGCAATGAGCTCGAGAAGGATTGGACAGTCTGT 2388  
QY 2238 GCATTTGTTGCCCTCCACTCTGATGGAATATGTTGTTTCTCCAGTCCCTCCACTCCAGAA 2297  
DB 2389 GTTTTGTGTTCTTACCCTAATGGAATATGTTGTTTCTCTCTTTCAGCAACTCCAGAA 2448  
QY 2298 TCTCTTTGAAAGAGCTATCCATGTCATTAAGTTTGGTGTATGAGGACAAAGTCTGAATGG 2357  
DB 2449 AACCTTCTCAAGAGGCTATCCATGTCATTAGTTTGGTTATGAGGATAAGTCAGATTGG 2508  
QY 2358 GGAATGAGATTGTTGGATCTATGATCTGTCACAGAAGATTCTTACTGATTCAG 2417  
DB 2509 GGAATGAGATTGATGATCTGTCACAGAAGATTCTTACTGATTCAG 2568  
QY 2418 ATGACGCAAGAGGCTGGGCTTCAGTCTATTGTCATGCCCAAGGCCAGCTTTCAAGGGA 2477  
DB 2569 ATGCATGCCGCTGGATGGGATCCATTACTGATGCCTAAGGTTCCAGCTTTCAAGGCT 2628  
QY 2478 TCTGCCCCATCAATCTTTCAGATCGTCTGAACCAAGTGGTGGGTGGGCTCTCGGTTCT 2537  
DB 2629 TCTGCTCTATCAATCTTTTCAGATCTCTGAACCAAGTCTGAGGTGGGCTTTAGGTTCA 2688  
QY 2538 GTTGAATCTTTTCAGCCGGCATTCGCCCTTATGTTGTTGCTACGAGGGGCCCTCAAG 2597  
DB 2689 GTTGAGATTCTTTCAGTCGGCATTTGCTCTATGATGTTGTTTACAAATGGGAGGCTA 2748  
QY 2598 TTCCGTGAGAGATTCCGTTTACATCAACACACACATTTACCACCTTCTCTCCCGCTT 2657  
DB 2749 TTTCTTTGAGGTTTGGTATGTGAACACACACATCTAGCCCTATCCCTCCATTCCTCT 2808  
QY 2658 CTAGTCTATGTTATGTCCTGCTATCTGCTCTCACTGGAAGTTTCATGCCAGAG 2717  
DB 2809 CTCATGTTATGTTACATTGTCAGCGGTGTGCTCTTCCACCAACCAAGTTTATTTCTCT 2868  
QY 2718 ATTGAACATTGGCCAGTATCTGGTTCATTGCTCTCTTCTTCAATTTTTCGCCACTGGT 2777  
DB 2869 ATTAGTAACATTGCAAGTATATGTTTCTGCTCTCTTCTCTCCATTTTCGCCAGGCT 2928  
QY 2778 ATCCCTGAGATGAGGTGGAGTGGTGTGGCAATTTAGCAGTGGTGGAGGAATGAACAGTTT 2837  
DB 2929 ATACTGAAATGAGGTGGAGTGGCGTAGGCATAGACGAATGGTGGAGAAACGACGAGTTT 2988  
QY 2838 TGGGTCAATTCGAGGTATCTCTGACATCTGTTTGGCGTCTTTCAGGGTCTTCTGAAGGTG 2897  
DB 2989 TGGGTCAATTCGAGGTATCTCTGACATCTGTTTGGCGTCTTTCAGGGTCTTCTGAAGGTG 3048

QY 2988 CTTGCCGCTATCGACACCAACTTCACTGTCAACCTCAAAAGCTAATGAAGAAAGGCGAC 2957  
DB 3049 CTTGCCGCTATGACACAAACTTCACATTTACCTCAAAAGCTTCAGATGAAGACGGAGAC 3108  
QY 2958 TTTGCTGAGCTCTACATGTTCAAGTGGACGCTTCTTCATCCCTCCGACGACCACTTTTG 3017  
DB 3109 TTTGCTGAGCTCTACTTGTCAATGGACAACACTTCTTGATTCGCCCAACGACGCTGCTC 3168  
QY 3018 ATCAATTAACATGTTGTTGCTGTTGGCACTCTTACGCCCATCAACAGTGGTTACCAA 3077  
DB 3169 ATTGTAATAGTTGGAGTTGTTGCAAGAGTCTCTTATGCTATCAACAGTGGATACAA 3228  
QY 3078 TCATGGGGCGCTCTTTGGGAAGCTCTTTGCTCTTCTGGTGGTGGTGGTGGTGGTGGTGG 3137  
DB 3229 TCATGGGGACCACTCTTTGTAAGTGTCTTTGCTCTCTGGGTAATGTTTCACTTGTAC 3288  
QY 3138 CATTTCTCAAGGCTTTATGGGACGCAAAACCGCACACCGACGATTTGTCATCTCTG 3197  
DB 3289 CCTTTCTCAAGGTTTATGGTGGTGGACAGAACCGGACTCTACCATTTGTTGTTGCTG 3348  
QY 3198 GCTGCTCTCTGCTTCTATCTTCTCTTCTGCTGTTGGTGGTGGTGGTGGTGGTGGTGG 3257  
DB 3349 TTTGCTCTCTGCTTCTATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3408  
QY 3258 GCTCTGCTGGCCCAATATCAAACTGTCGATCAACTGCTAGGAA 3305  
DB 3409 CGATCACTGGCCGGACATTTCTGAATGTCGAATCAACTGTTGAGAA 3456

RESULT 7  
AX030946  
LOCUS AX030946 3614 bp DNA linear PAT 20-SEP-2000  
DEFINITION Sequence 9 from Patent WO9800549.  
ACCESSION AX030946  
VERSION AX030946.1 GI:10278349  
KEYWORDS  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 3614)  
Williamson,R.E., Peng,L., Arioli,A. and Betzner,A.S.  
Manipulation of cellulose and/or beta -1,4-glucan  
Patent: WO 9800549-A 9 08-JAN-1998;  
WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI  
ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;  
COMMW SCIEN INT RES ORG (AU)  
Location/Qualifiers  
1. .3614  
/organism="Arabidopsis thaliana"  
/strain="COLUMBIA"  
/db\_xref="taxon:3702"  
/clones="ATH-B"  
217. .3414  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09627.1"  
/db\_xref="GI:10278350"  
/translation="MESEGETAGPKMKNIYPTQICSDNVGKTVGDREVFADICSF  
PVCRCYERKDNQSCPOCKTKYKRLKGSIPAIGDKDEGLADEGTVEFNYPQEK  
ISERMLGHLT RKGEMGPPQYDKVESHNLRLTSRQDTSGEFSFSAASPRLSVST  
TAGKRLPYSDVNQSPNRRIVDPVGLGNVAKERVDPGKWKOEKNTKPVSTOASER  
GGVIDASTDILADEALLNDEARQLLSRKYSIPSSRNIPYRMVMTLRLVLCFLHYR  
ITNPVNAFLWLVSVICEIWFALSWILDQFPKFWPVNRETYLDRLAURLDREGEPSQ  
LAAVDIFVTDVPLKEPLVTANTVLSIADYDPVKVSCVSDGGAAMLISFESLAET  
SEFAKWPKCYKYSIEBPAPWYFAAKIDYLDKDVQTSFVKDRRAKMKREYEEFKIRI  
NALYSKALKCEGWVMDGTPWPNNTGDHPGMIOVFLQNGGLDAGNELPLRVV  
SREKRPGFQHHKAGAMNALYVSALVTNGPFIILNDCDHVINNSKALREAMCFMDP  
NLKQVCYVQFPORPDGIDKNDRYANRNTVFYFLLNRLGLDGLGIVPYGTCGFLNRTA  
LYTEPPIKVKHPRKPSLLSLCGGSRKNSKAKSKDSKRSRDTSTVPVFNLDLIE  
EGVEGAGFDDEKALLMSQLEKFRGQSAVFVASTLMENGVPVPSATPEFLKPAIIV  
ISCYEDKSDWGMGMEIGWYISVTEIDLTGFKMHARGWRSIYCMFKLPFAKGSAPINLS

FEATURES  
source  
CDS

DRLNQVLRNALGSEVILFPRSHCDIWIYNGRLKFLERFAYVNTIPIPIPLMTCY  
LLAVCLTFNQFIPOISNIASINFLSFIATGDIEMRWSGVGDENWRNEQFVVI  
GGVSAHLFVAFVQILKLVLAGIDNFTVTSKASDEGDFAELYLFKWTLLIPPTLLI  
VNLVGVVAGVSVAINSGYWGMLFKLFFAFWVIVHLYPFLKGLMGRNRTPIVVV  
WSVLLASIESLLWVRIDPFTSRVTGPDILLECGINC.  
BASE COUNT 963 a 717 c 863 g 1071 t  
ORIGIN

Query Match 46.2%; Score 1673.6; DB 6; Length 3614;  
Best Local Similarity 71.4%; Pred. No. 0;  
Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;

QY	30	TCGGTGGCGTGTGATCGCTCTCGCGGCCCATGGACGGCGACGGCGGACGCCCTGAAATCC	89
DB	190	TCGGTGTGTGAAGCAACTAAGTAGCAAAATGGAATCCGAAGGAAACCCGGGGAACCG	249
QY	90	GGGAGGACGGCGCGGGGACGTGTGCCAGATCTGCCCGACGGCCTGGGCACCACTTG	149
DB	250	ATGAAGAACATTTGTCGCGAGACTTGCAGATCTGTAGTGACAATGTTGGCAAGACTGTT	309
QY	150	GACGGGACGCTCTTCACCGCTGCGAGCTGTGCCGCTTCCCGGTCTGCCGCCCTGCTAC	209
DB	310	GATGGAGATCGTTTGTGGCTTGTGATATTTGTTCATCCAGTTTGTGCGCCTTGCTAC	369
QY	210	GACACAGCGCAAGGAGGACCCAGGCGCTGCTCCAGTGCAGAACCAAGTACAAGCGC	269
DB	370	GAGTATGAGAGGAAGATGGAAATCAATCTGTCTCAGTCAAAACCAAGAGG	429
QY	270	CACAGAGGAGCCACGATCCCGGGGAGGAGGACGACACTGATGCCGATGATGGT	329
DB	430	CTCAAAGGTAGTCTCTATTCCTGTGTGATAAAGAGAGGATGGCTTAGCTGATGAAGT	489
QY	330	A---GTGACTTCAACTACCTGCTGGCACTGAGACCAAGAGCAGAAATTTGCTGAC	386
DB	490	ACTGTTGAGTTCACTACCT-----CAGAAGGAGAAATTTCAAG	531
QY	387	AGATGGCAGCTGGCGCATGACACCGGGGCGAGTGGCAATGTGGCCACCCCAAGTAT	446
DB	532	CGATGCTTGGTGGCATCTTACTCTGTGGGAAGGAGAGAAATGGGGAAACCCAGTAT	591
QY	447	GACATGGCGAGATCGGCTCTCCAGTAGTACAGTGGAGAGATCCCTAGGGATACGTC	506
DB	592	GATAAGAGG-----TCTCTCAATCATCTTCTCTGCTCAGGACGACAGCAAG--	640
QY	507	CCTTCAGTCAACCAACGCGATGTGAGGAGAAATCCCTGGAGCTTCGCTGATCATCAC	566
DB	641	-----ATACTTCAGGAGAGTTTCTGCTGCCCTCACTGAACGCTC	681
QY	567	ATGATGTCCCTACGGGGAACATCAGACAGCTGCTCGGTTTCCCT---ATGTAATCAT	623
DB	682	TCTGTATCTTACTATCTCGGTGGGAAAGCGCCTTCCCTATTCATCAGATGTCAATCAA	741
QY	624	TCACCAAAATCCGTCAGAGGAGTCTCCGCGCATTTGGGAATGTTGCCGTGGAAGAGAGA	683
DB	742	TCACCAATAGAAGAGTGTGGATCTGTGTGGACTCGGGAATGATGCTTGGAAAGAGAGA	801
QY	684	GTTGATGGCTGGAATAATGAAGCAGGACAAAGGTGGCATGCCATGACTAATGGCAACGC	743
DB	802	GTTGATGGCTGGAATAATGAAGCAGGACAAAGTACTGTGCTC-----TGTCACGACG	852
QY	744	ATTGCTCCCTCTGAAGGTGGGAGCTACTGACATCGATGCACTACTGAATACAACATG	803
DB	853	CAGGCTGCTTCTGAAAGAGGTGGAG---TAGATATTGTGCCAGCACAGATATCTAGCA	909
QY	804	GAAGACGCTTTACTGAATGATGAATCGCCAGCCTCTATCTAGAAAAGTCCCATGCT	863
DB	910	GATGAGGCTCTGCTGAATGACGAAGCGGAGCGAGCTTCTGTCAAGGAAAGTTTCAATTCCT	969
QY	864	TCCTCAAAATAAATCCCTACGAATGGTCAITGTTCTCGGGTGGTGTGTTCTTAAGCATC	923
DB	970	TCATCAGCGATCAATCCTTACAGATGGTATTATATCTCGGGCTTGTATCCTTTCTC	1029
QY	924	TTCTTGCACTACCGTCTCAAAATCCTGTGGTGAATGCAATACCCACTGGGCTTTATCT	983

DB	1030	TTCTTGCATTACCGTATAAACA	1089
QY	984	GTATATGTGAGATTTGGTTG	1043
DB	1090	GTGATATGTGAGATCTGGTTG	1149
QY	1044	CCATCAACCGGGAGACCTACG	1103
DB	1150	CTGTGAACCGTGAACCTACG	1209
QY	1104	CCGTCTCAGTTGGCTGCTGTTG	1163
DB	1210	CCATCAGATGTAGTCTGTTG	1269
QY	1164	CCATCTCTCACTGCCAACACTG	1223
DB	1270	CCCTTGTGACAGCAACACAG	1329
QY	1224	GTCTCTTGTCTATGTATCTGATG	1283
DB	1330	GTGCTCTGTTATGTTCTGATG	1389
QY	1284	ACTTCAGAGTTTGTAGAAAT	1343
DB	1390	ACATCAGAGTTTGTCTGTAAT	1449
QY	1344	GCCTCCGAGTTTACTTTTGGC	1403
DB	1450	GCACCAAGATTTGGTACTTTGCTG	1509
QY	1404	TTTGTAAAGACCGCCGGGCA	1463
DB	1510	TTTGTCAAAGATCGTAGAGCT	1569
QY	1464	GCCTAGTTTCTTAAGCATTTGA	1523
DB	1570	GCCTTTTCCAAAGCCCTA	1629
QY	1524	CCATGCCAGGAACATACCA	1583
DB	1630	CCGTGGCCTGGAATAATACAG	1689
QY	1584	AGTGGTGGCCTTGATCTGAGG	1643
DB	1690	AATGGTGGACTTGTGTCAGAGG	1749
QY	1644	AAGGCTCGTGGTTCAGCACG	1703
DB	1750	AAGCGACCGAGTATCCAGCAC	1809
QY	1704	TCAGCTGCTCTTACTAATGAGC	1763
DB	1810	TCAGCAGTTCTTACCATTGGAC	1869
QY	1764	AACAGCAAGGCTCTCCGAGAGG	1823
DB	1870	AACAGCAAGCCTTAAGAGAGG	1929
QY	1824	GTCTGTTATGTGAGTTTCCACG	1883
DB	1930	GTGTTGTTATGTTCACTTCCACG	1989
QY	1884	AACAGCAACTGCTCTTTTTCG	1943
DB	1990	AATCGTAATACCGTGTCTTTG	2049
QY	1944	GTGTTATGTGGGAACCTGTTG	2003
DB	2050	GTATATGTGCGAATCGATG	2109
QY	2004	ATTAAGG-----CGAAGAAGC	2057

Db 2110 ATAAAGTAACACACAGAGCCAAAGTCTTTTATCTAAGCTCTGTGGTGGATCAAGAAAG 2169  
QY 2058 GCAAGCAAGTCAAGAAAGAGAGCTCAGATAGAAAGAAAGTGAAGAGAGTGTGGACAGT 2117  
Db 2170 AAGAATTCACAAAGCTAAGAAAGAGTCCGACAAAGAAATCAGGCGAGCATACTGACTCA 2229  
QY 2118 TCTGTTCCAGATTAATCTCGAAGACATAGAGGAGGTGTGAAGGTGCTGGGTTCAT 2177  
Db 2230 ACTGTTCCGTGATTAACCTCGATGACATAGAGAGGAGGTGAAGGTGCTGGGTTCAT 2289  
QY 2178 GATGAGAAATCAGTTCTCATGCTCAAAATGAGCTTAGAGAGAGATTTGGCCAGTCAGCA 2237  
Db 2290 GATGAAAGGCGCTCTTAATGTCGAATGAGCTGGAGAGCGATTTGGACAGTCTGCT 2349  
QY 2238 GCAATTTGTCCTCCACCTCTGATGAATATAGGTGTGTTCTTCATGCTCTCCACTCCAGAA 2297  
Db 2350 GTTTTGTGTTCTTACCTAATGAAATAGTGTGTTCTCTCTTCAGCAACTCCAGAA 2409  
QY 2298 TCTCTTTTGAAGAGAGCTATCCATCTCATAGTCTTGCTGATGAGCAAGTCTCAATGG 2357  
Db 2410 AACTTCTCAAAGAGGCTATCCATGTCATAGTTGTTGTTAGGAGTAAGTCAGATTGG 2469  
QY 2358 GGAATCAGATTTGGTTGGATCTATGATCTGTACAGAGATATTTCTTACTGATTCAAG 2417  
Db 2470 GGAATGGAGATTGGATGGATCTATGTTCTGTGACAGAGATATTTCTGACTGGGTCAA 2529  
QY 2418 ATGACGGAAGAGGCTGGGTTTCACTATATGCTGATGCCAAGCGCCAGCTTTCAAGGA 2477  
Db 2530 ATGCATGCCGCTGGATGGGATGTCATTTACTGATGCCCTAAGCTTCCAGCTTTCAAGGT 2589  
QY 2478 TCTGCCCATCAATCTTTTACATGCTGTGAACCAAGTCTCGGTGGCTCTCGTTCT 2537  
Db 2590 TCTGCTCCATCAATCTTTTACATGCTGTGAACCAAGTCTGAGTGGGCTTTAGTTCA 2649  
QY 2538 GTTGAATCTTTTTCAGCGGATTTGCCCTTATGATGCTACGAGGCGCCTCAAG 2597  
Db 2650 GTTGAGATCTCTTCACTCGGATGTTCTATATGATGTTTACAATGGGAGGATAAA 2709  
QY 2598 TTCTCGAGAGATTGCTTTACATCAACACACCAATTTACCCACTAACCTCTCTCCGCTT 2657  
Db 2710 TTTCTTGAGAGGTTTTCGATGTGACACACCACTACCCCTATCACCTTCAATCTCTT 2769  
QY 2658 CTAGTCTATGATATGCTGCTATCTCTGCTCACTGGAAGTTTATCATGTCAGAG 2717  
Db 2770 CTATGATTTGATGTTGCTAGCGGTTGCTCTTTCACCAACAGGTATTTATCTCTCAG 2829  
QY 2718 ATTAGCACTTGGCCAGTATCTGTTTCAATGCGCTCTTCCCTTCAATTTTCCGCACTGGT 2777  
Db 2830 ATTAGTAACATTGCAAGTATATGGTTTCTGCTCTCTTCTCTCCATTTTCCGCACTGGT 2889  
QY 2778 ATCCTTGAGATGAGTGGAGTGGTGTGTCATGAGAGTGTGGAGGAATCAACAGTTC 2837  
Db 2890 ATACTAGAATGAGTGGAGTGGCTGAGCATAGAGAAATGTTGAGAAACGAGAGTTT 2949  
QY 2838 TGGCTCATTTGAGGATATCTCTGCATCTGTTTGGCGTCTTTCAGGCTTCTTGAAGTGT 2897  
Db 2950 TGGCTCATTTGAGGATATCTCTGCATCTGTTTGGCGTCTTTCAGGCTTCTTGAAGTGT 3009  
QY 2898 CTTGCGGATATGACACCACTTCACTGTCACCTCAAGGCTTAAGCAAGAGCGAC 2957  
Db 3010 CTTGCGGATATGACACCACTTCACTGTCACCTCAAGGCTTAAGCAAGAGCGAC 3069  
QY 2958 TTTGCTGAGCTTACATGTTCAAGTGGAGAGGCTTCTATCCCTCCGAGCACTTTTG 3017  
Db 3070 TTTGCTGAGCTTACATGTTCAAAATGGACAACACTTCTGATTCGCGCAACAGCTGCTC 3129  
QY 3018 ATCATTAAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3077  
Db 3130 ATTGTAACACTTAGTGGAGTGTGTCAGAGAGTCTTATGCTATCAACAGTGGATACCAA 3189  
QY 3078 TCATGGGGCGGCTCTTTTGGAGAGCTCTTCTTTGCGCTCTGCGCTGATTTGCTCACTATAC 3137  
Db 3190 TCATGGGGCACTCTTTTGGTAAGTGTCTTTTGGCTTCTGCTGCTGCTGCTGCTGCTGCT 3249

QY 3138 CCATTCTCAAGGCTTTATGGGCGAGGCAAAACCGCACCGGAGATTGTTCATCGTCTGG 3197  
Db 3250 CTTTCTCAAGGCTTTATGGGCTGAGAGACCGGACTCTTACCATTTGTTGGTCTGG 3309  
QY 3198 GCTGCTCTCTCGCTTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3257  
Db 3310 TCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3369  
QY 3258 CGTCTGCTGCGCCAAATATCAAAACCTGTGGCATCAACTGCTAGGAA 3305  
Db 3370 CGAGTCACTGCGCGGACATTTGGAATGTGGAATCAACTGTTGAGAA 3417

RESULT 8  
AX030960  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
thale cress.  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 3614)  
Williamson, R.E., Peng, L., Arioli, A. and Betzner, A.S.  
Manipulation of cellulose and/or beta -1,4-glucan  
Patent: WO 9800549-A 08-JAN-1998;  
WILLIAMSON RICHARD EDWARD (AU) ; PENG LIANGCAI (AU) ; ARIOLI  
ANTONIO (AU) ; UNIV AUSTRALIAN (AU) ; BETZNER ANDREAS STEFAN (AU) ;  
COMWU SCIENT IND RES ORG (AU)  
FEATURES  
Location/Qualifiers  
1..3614  
/organism="Arabidopsis thaliana"  
/strain="COLUMBIA"  
/db\_xref="taxon:3702"  
/clone="ATH-B"  
217..3414  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC09633.1"  
/db\_xref="GI:10278364"  
CDS  
translation="MESEGTACKPMKNIVPOTCIQSDNVGKTVGDREVCADICSF  
PVCRCPEYERKDGNSQCPCKTRYKLGKSPALPGDKDDEGLADEGTFEYFNYPQKEK  
ISERMLGHLTRGGEEMPEPQFDKEVSHNHLRLSRQDTSFEFSAERLSVST  
IAGGKRLPYSSDVNQSPNRRIVDPGLGNVAMKERVDMKMKOENKNTGPVSTQASER  
ITNPVPNAFALWLVCEIWFALSIWLDQFPKFPVNRETYLDRLALRYDREGPSQ  
LAAVDLFTVSDPLKPPPLVTANTVLSILADYDVPDKVSCYSDSDDGAAMLSPESLAEI  
SEFARKVPECKKYSIEPRAPENYFAKIDYLDKQVTSFKDRAKREYEFKIRI  
NALYSKALCKPESGWMDGTPWPNNTGDHPGMIQVFLQNGGLDAEGNELPRLVYV  
SRKRPGFQHHKAGAMNALRVSAVLNTGPFILNLDCHYINNSKALREAMCFLMDP  
NLGKQCVYQFPQFDIDKNDRYANRNTVFDINLRGLDGIQGPVYGTGCVFNRTA  
LYGVEPIKVKHKKPSLLSKLGGSKRKNKSKAKESDKKSGRHTDSTVGFVNLDDIE  
EGVEGAGDEKALLMSLEKRFQGVSAFVASTLMENGVPVSPATPENFLKATHV  
ISCGYEDKSDWMEICWIVGVTEDILTGPKHARGWRSIYCMKLPAPKGSAPINLS  
DRNLQVLRWALGSVEILFSRHCPIWYNGYRGLKFLERFAVNTTIPISILMNTCT  
LLAVSLFTNQFIIPQISNIASIMWLSLFIATGILEMWSGVDEWMNRNQFWI  
GGVSAHLFAVGILKVIAGIDNTFTVTSKDEGDAELYLFKWTLLIPIPTLLI  
VNLGVVAGVSVAINSQYQSWGPLFGKLPFAFWIVHLYPLFLKGLMGRQNRTPTVV  
WSVLASIFSLNVRIDPFTSRVTGPDILECGINC"

BASE COUNT 963 a 717 c 863 g 1071 t  
ORIGIN

Query Match 46.2%; Score 1673.6; DB 8; Length 3614;  
Best Local Similarity 71.4%; Pred. No. 0;  
Matches 2347; Conservative 0; Mismatches 869; Indels 72; Gaps 8;

QY 30 TCGGTGCGGTTGGATCGCTCTGCGGCCCATGAGCGGACGCGGAGCCCTGAAGTCC 89  
Db 190 TCGGTGTTGGAAGCAACTAAGTGACAATGGAATCCGAAGGAGAAACCGCGGAAAGCCG 249

QY 90 GGGAGGCGCGGGGACGTGTGCCAGATCTGGCCGACGCGCTGGGCGACACCGTTG 149  
DB 250 ATGAAGAACAATGTTCCGACAGCTGCGAGATCTGTAGTGACAATGTGGCAACAGCTT 309  
QY 150 GACGCGACGCTCTTACCGCTGGGAGCTGTCCGGCTTCCGGTCTGCGCGCCCTGCTAC 209  
DB 310 GATGGAGATCGTTTGTGGCTGTGTATTTGTTCAATTCGCCAGTTTGTGCGGCCCTTGCTAC 369  
QY 210 GAGCAGCGACGAGGAGGCGACCGGCTGCTCAGTCCAGACCAAGTACAAGCGC 269  
DB 370 GAGTATGAGAGGAAGATGGGAATCAATCTTGTCTCAGTGCAAAACCCAGATACAAGAG 429  
QY 270 CACAGAGGCGCCAGCGATCCGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 329  
DB 430 CTCGAAGGTAGTCTGCTATTCCTTGGGTATTAAGACGAGGAGTGGCTTAGCTGATGAAGGT 489  
QY 330 A---GTGACTTCAACTACCTGCATCTGGCACTGAGGACCAAGAGAGAGATGCTGAC 386  
DB 490 ACTGTTGAGTTCAACTACCT-----CAGAAGGAGAAAATTCAGAG 531  
QY 387 AGGATGCCAGCTGGCGCATCAACCGGGGGCAGTGGCAATGTTGGCCACCCCAAGTAT 446  
DB 532 CGGATGCTTGGTGGCATCTTACTCGTGGGAAAGGAGGAGAAATGGGGGAACCCCAAGTAT 591  
QY 447 GACAGTGGCGAGATCGGCGCTCTCCAAGTATGACAGTGGAGAGATCCCTAGGGGATACGTC 506  
DB 592 GATRAAGAGG-----TCTCTCACAATCATCTCTCGCTCACGAGCAGACAAG-- 640  
QY 507 CCTTCAGTCAACACGCGCATGTGAGGAGAAATCCCTGGAGCTTCGCCGTGATCAAC 566  
DB 641 -----ATACTTCAGGAGGTTTCTGCTGCGCTCACCTGAAACGCGCTC 681  
QY 567 ATGATGCTCCCTACGGGGAACATCAGCAGAGCTGCTCGCTTTCCT---ATGTCGAATCAT 623  
DB 682 TCTGTATCTTACTATCGCTGGGGAAAGCGCCTTCCCTATTATCAGATGTCATCAACAA 741  
QY 624 TCACCAAAATCGGTCAAGGAGTTCCTCGGCGAGTATGCGGAATGTGGCTGGGAAAGAGAGA 683  
DB 742 TCACCAAAATAGAAAGATTTGTGGACTCTGTTGGACTCGGGAATGTAGCTTGGGAGGAGAGA 801  
QY 684 GTTATGCTGCGAAATCAAGCAGCAGCAAGGTCGATTCCTCATGACTAATGCGCAGCAGC 743  
DB 802 GTTGATGCTGGAATGAACAGAGAGAGAAATACTGGTCC-----TGTGAGCAGCAG 852  
QY 744 ATTGCTCCCTCTGAAGTCTGGGCGAGCTACTGACATGATGATGATCTACTGATATCAACATG 803  
DB 853 CAGGCTGCTTCTGAAAGAGGTGGAG---TAGATATATGATGCGCAGCAGATATCTAGCA 909  
QY 804 GAAGACGCTTACTGAATGATGAACCTGCGCAGCGCTCTAICTAGAAAAGTCCCATTTGCT 863  
DB 910 GATGAGGCTCTGCTGAATGAGCAAGCGGAGGAGCTCTGCTCAAGGAAAGTTTCAATTTCT 969  
QY 864 TCCCTCAAAATAAATCCCTACAGAATGCTGCTGCTGCGGTTGGTTTCTTAAGCATC 923  
DB 970 TCATCAGGATCAATCTTACAGAATGTTATATGCTGCGGCTGTATCTCTTGTCTC 1029  
QY 924 TTCCTGCACTACCGTCTCAGAAATTCCTGTCGCTGATGCAATACCACCTGTCGCTTTATCT 983  
DB 1030 TTCCTGCACTACCGTCTCAGAAATTCCTGTCGCTGATGCAATACCACCTGTCGCTTTATCT 1089  
QY 984 GTTATATGAGATTTGGTTTGGCTTTATCTCTGATGATGATGATGATGATGATGATGAT 1043  
DB 1090 GTGATATGAGATCTGTTTGGCTTATCTCTGATTTGGATGATGATGATGATGATGATGAT 1149  
QY 1044 CCAATCAACCGGAGACCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1103  
DB 1150 CCGTGAACCGTGAACCTACTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1209  
QY 1104 CCGTCTCAGTTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1163  
DB 1210 CCAATCAGATGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1269

QY 1164 CCTATCGTCACTGCCAACACTG 1164  
DB 1270 CCCCCTGTGACAGCCACACAG 1329  
QY 1224 GTCTCTGCTATGATGATGATG 1283  
DB 1330 GTGCTCTGTTATGTTTCTGATG 1389  
QY 1284 ACTTCAGAGTTTCTAGGAAT 1343  
DB 1390 ACATCAGAGTTTCTGCTGAT 1449  
QY 1344 GCTCCCGAGTTTACTTTTCTG 1403  
DB 1450 GCACAGATGTTTACTTTGCTG 1509  
QY 1404 TTTGTTAAAGCCCGCGGCGCT 1463  
DB 1510 TTTGTTCAAGATGTTAGAGCT 1569  
QY 1464 GCCTTACTTCTAAGCATTTG 1523  
DB 1570 GCACTTGTTCRAAGCCCTTA 1629  
QY 1524 CCATGGCCAGGAACATATCA 1583  
DB 1630 CCGTGGCTGGAATAATACAG 1689  
QY 1584 AGTGGTGGCTTGTATGATGAG 1643  
DB 1690 AATGGTGGATTTGATGAGAG 1749  
QY 1644 AAGCGTCTCGGTTTCCAGCAC 1703  
DB 1750 AAGCGACAGATTCACAGCAC 1809  
QY 1704 TCAGCTGCTCTTACTAATGAG 1763  
DB 1810 TCAGCAGTTTCTACCAATGAG 1869  
QY 1764 AACAGCAAGCTGTCCGAGAG 1823  
DB 1870 AACAGCAAGCTTAAAGAGAG 1929  
QY 1824 GTCTGTTATGCGAGTTTCCAC 1883  
DB 1930 GTTGTGTTATGTTCAAGTTCC 1989  
QY 1884 AACAGCAACACTGTCTTTTGT 1943  
DB 1990 AATCGTAAATCCGTTGTTCT 2049  
QY 1944 GTTATGTTGGAACTGTTGTTG 2003  
DB 2050 GTATATTCGGAACCTGATGAT 2109  
QY 2004 ATTAAGG-----CGAAGAG 2057  
DB 2110 ATAAAGTAAACACAAAGAGAG 2169  
QY 2058 GCAGCAAGTCAAGAAAGAGAG 2117  
DB 2170 AAGAATTCCAAGCTAAGAAAG 2229  
QY 2118 TCTGTTCCAGTATTCATCTCG 2177  
DB 2230 ACTGTTCTCTTATTCACACCT 2289  
QY 2178 GATGAGAAATCAGTTCATGAT 2237  
DB 2290 GATGAAAGGCGCTCTTAATG 2349  
QY 2238 GCATTTGTTGCTTCCACTCTGA 2297





QY 173 CGACGCTGCGCTTCCCGTCTCCGCGCTCTAGAGCAGGCGCAAGGCGCAC 232  
DB 358 CAATGAGTGTGCCCTTCCCTGTCTGCGCGCTTCTGTATGATGAGCGCGCAAGGAGGAA 417  
QY 233 CCAGGCGTGCCTCCAGTGGCAAGACCAAGTACAAAGCGCCACAGAGGAGCCAGCGATCCG 292  
DB 418 CCAATGCTGCCCTCAGTGCAGAACTAGATACAGAGACAGAAAGGTAGCCCTCGAGTTCA 477  
QY 293 CGGGGAGGAAGCGGACACACTGATGCCGATGATGGTAGTGACTTCAACTACCCCTGCATC 352  
DB 478 TGGTGATGATGAGGAGGAAGATTTGATGACCTGGCAATGAATCAACTATAGAAGG 537  
QY 353 TGGCACTGAGGACCAAGACAGAGAGATTCGTGACAGGATCGCGAGCTGGGCGCATGACAC 412  
DB 538 CAATGGGAAGGGCCAGAGTGGCAGCTT-----CAAGGAGAT 574  
QY 413 CGGGGCGAGTGGCAATTTGGCCACCCCAAGTATGACAGTGGGAGATCGGCCTCTCCAA 472  
DB 575 GACGCTGATGTCTTCATCTGCTGCCATGACCCACACCATCGGATTCACGCGCTTACA 634  
QY 473 GTATGACAGTGGAGATCCCTAGGGGATACGTCCTTCAGTCACCAACAGCCAGATGTC 532  
DB 635 AGTGGACA-----ACAGATATC 651  
QY 533 AGGAGAAATCCCTGGAGCTTCGCTGATCATCATGATGTCGCCCTACGGGGAACATCAG 592  
DB 652 TGGAGAGATCCCTGATGATCCCTGACCGTCAATTCATCCGCGAGTCCAAACATCGAGCTA 711  
QY 593 CAGACGTGCTCGCTTCCCTATGTGAATCATTCACCAATTCCTCAAGGGAGTT---CTC 649  
DB 712 TGTGTATCAAGGCTCCAGTTCCTCTGAGGATTTGGACCCCTCGAAGGACTTGAATTC 771  
QY 650 CGGCACTATTGGGAATTTGCCCTGGAAAGAGAGAGTTGATGGCTGGAAATGAAGCAGGA 709  
DB 772 CTATGGCTTTAATGTTGATGCTGGAAGGAAGAGTTGAGAGCTGGAGGTTAAACAGGA 831  
QY 710 CAAGGCTGGGATTCCTGACTAATGGGACAGCATTTGCTCCCTCTGAAGGTGGGCGAGC 769  
DB 832 CAAAAATATGTCGAAGTGAATA-----TAAATATCCAGAGGC 870  
QY 770 TACTGACATCGATCTACTGAATACACATGGAGAGCGCTTTACTGAATGATGAAC 829  
DB 871 TAGAGAGACATGGAGGAGCTGGCTCAAAATGGAGAGATATGCAAAATGGTTGATGATGC 930  
QY 830 TCGCCAGCCTCTATCTAGAAAAGTCCCAATTCCTCCAAATAAAATCCCTACAGAAAT 889  
DB 931 ACCCTACCTTTGAGCGGCAATTGCGCAATTTCTCAACACGCTCAACCTTTACCGGAT 990  
QY 890 GGTCAATGTTTCGGTGTGTTCTAAGCATCTTCCTGCACTACCGTCTCACAATCC 949  
DB 991 AGTAATCAATTCCTCGTCTTATCATCTCTGCTCTTCTTCCCAATATCGTATCAGTCATCC 1050  
QY 950 TGTGCTTAATGATACCCACTGCGCTTTATCTGTATATGTGAGATTTGGTTTGCCTTT 1009  
DB 1051 AGTGCTAATGCTATGAGATTTGGGTAGTATCTGTATCTGTGAGGCTGTTGGCTTT 1110  
QY 1010 ATCCTGGATCTGGATCAGTTCCCGAAGTGGTTTCCAAATCAACCGGAGACCTACCTTGA 1069  
DB 1111 GTCTGCTCTAGATCAGTTCCCAAAATGGTATCCAAATCAACCGTGAACATATCTCGA 1170  
QY 1070 TAGACTGGCTTTAAGGTATGACCGAAGGTGAACCGTCTCAGTTGGCTGCTTTGACAT 1129  
DB 1171 CAGGCTTGCATTTAGGTATGATAGAGAGGAGAGCCATCAGACTGGCTGCCATTTGATGT 1230  
QY 1130 ATTTGTGCTAGTACGACCCCTTGAAGGAGCCACTATCGTCACCTGCCAACACTGTCT 1189  
DB 1231 CTTTGTGCTAGTGGATTCATTTGAGGAACTCCACTGATCAGAGCCAACTGTTTT 1290  
QY 1190 ATCCATCTTGTGTTGATTTATCCCTGGACAAAGGCTCTTGTCTATGCTATCTGATCAGCG 1249  
DB 1291 GTCCATCTTGTGTTGATTTACCTCTGTGACAAAGTGTCTATGTTCTGATGATGG 1350

QY 1250 AGCTTCAATGCTGACTTTTGAC 1309  
DB 1351 CTGAGTATGCTGACTTTTGAC 1410  
QY 1310 ACCATTGTGGAAGATGATGAC 1369  
DB 1411 TCCTTTTGTGAAGACACAA 1470  
QY 1370 AATTGATTACCTGAAGACAA 1429  
DB 1471 AATAGATTACTTGAAGCAAA 1530  
QY 1430 GAGAGATATGAAGATTTAA 1489  
DB 1531 GAGAGATATGAAGATTTCAA 1590  
QY 1490 CCCGAGGAAGGATGGATCAT 1549  
DB 1591 GCCTGAAGAGGGTGGACCAT 1650  
QY 1550 TCATCTTGGAAATGATTCAGGT 1609  
DB 1651 CCATCTTGGCATGATTCAGGT 1710  
QY 1610 TGAGCTCCCCCGTTTGTATTT 1669  
DB 1711 TGAATTACCAGCTCTTGTCT 1770  
QY 1670 GAAGGCTGGTGCATGAATGCC 1729  
DB 1771 GAAGGCTGGTGCATGAATGCC 1830  
QY 1730 CATGTTGAATCTTGTATGAT 1789  
DB 1831 TCTTCTCAATGTGATTTGAC 1890  
QY 1790 GTCTTCTTAATGGATCCAAAG 1849  
DB 1891 GTCTTCTCATGATGGATCCAG 1950  
QY 1850 GTTGTGATGGATTCATAGGAAT 1909  
DB 1951 ATTGTGATGGATTCATAGTCAC 2010  
QY 1910 TAACTTTCAGGGGCTTGACGGC 1969  
DB 2011 CAACATGAAGGTCTAGATGCG 2070  
QY 1970 CAACAGAACGCTATCTATGCT 2026  
DB 2071 CAATAGGCGAGCTTTGATGGA 2130  
QY 2027 CTTTGGCATCACTATGTGGGGC 2086  
DB 2131 CATTGTTTGAAGCTGTGCTGT 2190  
QY 2087 TAAGAAAAAGTTCGAACACAT 2146  
DB 2191 TCAAGCCGATTTATGAAGAG 2250  
QY 2147 AGAGGAGGTGTTGAAGGCTCT 2206  
DB 2251 CGAGGAGGTATTCAGAGT--- 2304  
QY 2207 GACCTTAGACAAGAGATTTGGC 2266  
DB 2305 GAAATGGAGAAAGCTTTGGT 2364  
QY 2267 TGGTGTGTTTCCCTCAGTCCCT 2326  
DB 2365 AGGTGGCATACCACCTTTCAACA 2424  
QY 2327 AAGTTGTGGCTATGAGGACAA 2386







Db 546 TATTGATCCACGGCAACCTGTCCTGTAAAGAAATCGTGGACCCGTCACAAAGACTT----- 599  
Qy 631 ATCCGTCAGGGAGTTCTCCGGCAGATATGGGAATGTTCCCTGGAAAGAGAGACTGTATG 690  
Db 600 -----GAACCTATTATGGCGTTGGTAATGTTGACCTGGAAAGAAAGAGTTGAAG 646  
Qy 691 GCTGGAATAATGAAGCAGGACAAAGGGTCGATTCCCATGACTAATGGGACAAAGCATGTCTC 750  
Db 647 CTTGGAAGCTGAAGCAGGAGAAAATATGTTACAGATGAC-----TGATA 691  
Qy 751 CCTCTGAAGTTCGGGAGCTACTGACATCGATGCACTACTCAATACACATGGAAGACG 810  
Db 692 AATFACATGAAGGAAGAGGAGAGAAAT---TGAAGGAGTGGTTCCCAATGGCGAAGAAC 748  
Qy 811 CTTTACTGAATGAACACTCGCCAGCCCTTATCTAGAAAAGTCCCATTCCTCCCA 870  
Db 749 TCCAAATGGGTGATGATACACGCTCTCCCTATGAGTGGTGGCTATCCCATCTTCTC 808  
Qy 871 AAATAAATCCCTACAGAAATGTCATTGTTCTCGGGTTGTTCTTAAGCATCTTCCCTGC 930  
Db 809 GCCTAACCCCTTATCGGGTTGATTAATCTCCGGCTTATCATCTGTGTTCTTCTTCTGC 868  
Qy 931 ACTACCGTCTCAAAATCTCTGGTAAATGATGATACCATCGGTGTTTATCTGTATAT 990  
Db 869 AATATCGTCAAACTCACCCCTGTGAAAATGCATATCCTTTGTTGTTGACCTCGGTTATCT 928  
Qy 991 GTGAGATTGGTTGCTTTATCTCGGATCTGATCAGTCCGAGTCCGAGTGGTTTCCCAATCA 1050  
Db 929 GTGAGATCTGGTTGCAATTTCTTGGCTCTTGGATCAGTTTCCCAATGGTACCCCATTA 988  
Qy 1051 ACCGGGAGACCTACCTGTATGATAGACTGGCTTAAAGTATGACCGAAGAGTGAACCGTCTC 1110  
Db 989 ACAGGAGACCTATCTTACCGCTCTCGCTATAAGATATGATCGAGACGCTGAACCATCAC 1048  
Qy 1111 AGTTGGCTGTGTTGACATATTTGTCAGTACAGTGCACCCCTTGAAGGAGCCACCTATCG 1170  
Db 1049 AGCTCGTCTGTTGATGCTTTGTTAGTACAGTGACCCCATGAAAGAGCTCCCTTG 1108  
Qy 1171 TCACGTCCACACTGTGCTATCCATTTCTGTTGCTGTTGATATCCCTGGACAAAGTCTCTT 1230  
Db 1109 TTACAGCAACACAGTTCTCTCGATTTCTTCTGTGGAGTACCCGCTAGATAAAGTAGCCT 1168  
Qy 1231 GCTATGATCTGATGACGGAGCTCAATGCTGACTTTTGAACCATTTGCTGAGACTTTCAG 1290  
Db 1169 GTTATGTTTACAGATGATGTTTACGCTATGCTTACCTTTGAAATCCCTTCTGNAACCGTG 1228  
Qy 1291 AGTTTGTAGAAATGGGTACCATTTGTTGAAGAAGTATGACATTTGAACCCAGAGCTCCG 1350  
Db 1229 AGTTTGAAGAAATGGGTACCATTTTGAAGAAATTTCAACATTTGAACCTAGGGCCCTG 1288  
Qy 1351 AGTTTGTACTTTGCGCAAAATTTGATTACTTGAAGACAAAGTCCAGCTTTCATTGTGTA 1410  
Db 1289 AATTCTATTTGCGCAGACATAGATTACTTGAAGCAAGATCAACCGCTTTGTGTA 1348  
Qy 1411 AAGACCCCGGGCCATGAGAGAGAAATATGAAGATTTAAATACAGGATAAATGCCCTAG 1470  
Db 1349 AAGAGCAGCAGCTATGAGAGAGAGATATGAAGAGTTTAAAGTGAGGATAAATGCTCTTG 1408  
Qy 1471 TTTCTAAGGCATTGAAGAGTCCCGAGGAGGATGGATCATCAAGATGGCACACCATGGC 1530  
Db 1409 TTGCAAAAGCACAGAAAATCCCTGAAGAGGCTGGACATGACAGATGGTACTCCCTGGC 1468  
Qy 1531 CAGAAACAATACAGGGATCATCTCGGAATGATTCAGGTTTTCCTGGTCACAGTGGTG 1590  
Db 1469 CTGTAACAACACTAGAGATCATCTTGGAAATGATACAGGTGTTCTAGGCCATAGTGGGG 1528  
Qy 1591 GCCTTGTACTAGGGTATGAGCTCCCGCGTTTATGTTTATGTTCTGCTGCTGAAAGCGTC 1650  
Db 1529 GTCTGGATACCGATGGAATGAGCTGCTAGACTCATCTATGTTCTCTGTAAGAGCGGC 1588  
Qy 1651 CTGGGTTCCAGCACCACAGAGGCTGGTCCCATGAAAGCCCTTGTGTTGCTGCTCAGCTG 1710  
Db 1589 CTGGATTTCACACCAACCAAGGCTGGAGCTATGAATGCAATTGATCCGCTGATCTGCTG 1648

Qy 1711 TCCTTACTAATGCAACAATACAT 1770  
Db 1649 TTCCTTACCNAATGGAGCATATCA 1708  
Qy 1771 AGGCTGTCGGAAGCACTATGTC 1830  
Db 1709 AGGCTATTAAAGAAAGCTATGTC 1768  
Qy 1831 ATGTGCAAGTCCCAACAAGGTT 1890  
Db 1769 ATGTCCAGTTCCTCAAGCTTC 1828  
Qy 1891 ACAGTGTCTTTTTTGTATATTA 1950  
Db 1829 ATATAGCTCTTTTTCGATATTA 1888  
Qy 1951 TGGGAACCTGGTTGTCTTTCAA 2003  
Db 1889 TGGGTACTGGTTGTGTTTAA 1948  
Qy 2004 ATTAAGCGGAAGAACCGAGGTT 2061  
Db 1949 AAGAAGATTTAGAACCAAAATAT 2008  
Qy 2062 GCAAGTCAAGAAAGAAAGGAGCTC 2121  
Db 2009 AAAGTAGCAAGAAAGTAACTAC 2068  
Qy 2122 TTCAGTATTTCAATCTCGAAGAA 2181  
Db 2069 CTCCTACTTTTCAATATGAGGAA 2122  
Qy 2182 AGAAATCAGTCTCTCATGCTCAA 2241  
Db 2123 AGAGGCTATTCTAATGTTCCAG 2182  
Qy 2242 TTGTTGCCCTCCACTCTGATGGA 2301  
Db 2183 TTATTCGGGCAACCTTTCATGGA 2242  
Qy 2302 TTTTGAAGCAAGCTATCCATGTC 2361  
Db 2243 TTCTGAAGGAGCTATTTCATGTC 2302  
Qy 2362 CTGAGATTTGGTTGGATCTATGGA 2421  
Db 2303 AAGAGATTTGGTTGGATCTATGTC 2362  
Qy 2422 AGCAAGAGGCTGGGCTTCAGTCT 2481  
Db 2363 ATGCCCGGTTGGATATCGATC 2422  
Qy 2482 CCCCATCAATCTTTTCAGATCGTC 2541  
Db 2423 CACCAATCAATCTTCTTATCGTC 2482  
Qy 2542 AAATCTTTTTCAGCGGAGTTCG 2601  
Db 2483 AGATTTCTTTAGCAGACATGTC 2542  
Qy 2602 TGGAGAGATTCCTTTACATCAAC 2661  
Db 2543 TGGAGAGGATGCGTTATATCAAC 2602  
Qy 2662 TCTATTGTATATGCTGCTATCT 2721  
Db 2603 CGTATTGTATCTTCCCGCTTTT 2662  
Qy 2722 GCAAGTTCGCAAGTATCTGGTCA 2781  
Db 2663 GCAACTACCGAGTATTTGGTTCA 2722



BASE COUNT	950 a	745 c	852 g	1056 t
ORIGIN				
Query Match	37.0%	Score 1343.4	DB 8	Length 3603
Best Local Similarity	66.2%	Pred. No. 2.4e-254		
Matches 2098	Conservative 0	Mismatches 991	Indels 80	Gaps 2
QY	92	GAGGCACGGGGCCGGGACGCTGCCAGATCTGCGCGACGCGCTGGGCACACACGTTGA	151	
DB	93	GAAGAATAAGATGGCCAGATATGTCAGATCTGTGGTGATGATGTTGGACTCGCTGAAC	152	
QY	152	CGCGACGCTTTCACCGGCTGCGACGCTGCGCGTCCCGGCTTCGCCGCTGCGCGCCCTGCTACGA	211	
DB	153	TGGAGATGCTTTGTGCGGTGTAATGAATGTGCCTTCCTCTGTGTGCGGCTTGCTATGA	212	
QY	212	GCACGAGCCAGGAGGACCCAGGCGCTGCCTCCAGCTGCAGACGACAGCAACGATACAGCGCA	271	
DB	213	GTACGAGAGGAAGATGAATCTAGTGTGGCCCTCAATGCACACACTAGATTCACAGCACA	272	
QY	272	CAGAGGGACCCAGCGATCCCGGGGAGGAGCGACGACACTGATCGCGATGATGGTAG	331	
DB	273	CAGGGGAGCTCCCTGTTGTGAAGGATGAGAAATGAGGATGATGTTGATGATATCCAGAA	332	
QY	332	TGACTTCAACTACCTGTCATCTGGCACTGTGGCACTGAGGACGACAGAAAGATTGCTCAGAGGAT	391	
DB	333	TGACTTCAATTAGC-----CCAGGGAGCTTACCAAGCC	365	
QY	392	GGCAGCTGGCGGATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAG	451	
DB	366	GAGACACCAACGCCATGGCGAAGATTTTCTCTCTAGACATGAATCTCAACCAAT	425	
QY	452	TGGCGAGATCGGCCTCCCAAGTATGACAGTGGAGAGATCCCTAGGGGATACGTCCTTC	511	
DB	426	TCCTCTTCTACCCATGCCATACGGTTTCTGGAGAGATTGCGACGCCGCTATACAAATC	485	
QY	512	AGTCCACCAAG-CCAGATGTGAGGAGAAATCCCTGGAGCTTCGCTGATCATCATGA	570	
DB	486	TGTGCGAACTAGATCAGGTCCCTTTGGGTCTCTGACAGGAATGCTATTTCATCTCCATA	545	
QY	571	TGTCCTCTACGGGAAACATCAGCAGACGTGCTCGCTTTTCCCTATGTGAATCATTCACCAA	630	
DB	546	TATTGATCCAGCCAACTGCTCCCTGTGAAGATCGTGGCCCGCTCAAAAGACTT-----	599	
QY	631	ATCCGTCAAGGAGTTTCCCGCAGTATTGGGAATGTTGGCTGGAAAGAGAGAGTTGATG	690	
DB	600	-----GAACTCTATTGGCTTGTGTAATGTTGACTGAAAGAAAGAGTTTGAAG	646	
QY	691	GCTGGAAATGAAGCAGGACGAGGTGCGATTCCTCATGTGGAATCATTCACCAA	750	
DB	647	GCTGGAAGCTGAAGCAGGAGAAATAATGTTACAGATGAC-----TGGA	691	
QY	751	CTCTGAAGGTGGGCGAGTACTGACATCGATGCTACTGTAATACAACTGGAAGACG	810	
DB	692	AATACCATTGAAGGAAAGSAGAGAAAT---TGAAGGAGCTGTTCCATTCGCGAAGAAC	748	
QY	811	CTTTACTGATGATGAACCTGCCAGCTCTATCTAGAAAAGTCCCATTCCTTCCTCA	870	
DB	749	TCCAAATGGCTGATGATACAGCTTCTCTAGCTGCTGGTGCCTATCCCATCTTCTC	808	
QY	871	AAATAAATCCCTACAGAAATGGTCATGTTCTGGGGTGTGTTGTTCTAAGCATCTTCTCG	930	
DB	809	GCCTAACCCCTTATCGGTTGTGATTAATCTCCGCGCTTATCATCTTGTGTTTCTTCTCG	868	
QY	931	ACTACCGTCTCAAAATCTGTGCGTAATGATACACCACTGTGGCTTTTATCTGTATAT	990	

Db 869 AATATCGTAACTCACCTGCGTGAATAATGCATATCCTTTGGTGTGACCTCGGTTATCT 928  
QY 991 GTGAGATTTGGTTGCTTTATCCTGGATCTGGATCAGTTCCTCGAAGTGGTTTCCAATCA 1050  
Db 929 GTGAGATCTGGTTGCAATTTCTTGGCTTCTTGATCAGTTCCTCCAAATGGTACCCCAITA 988  
QY 1051 ACCGGGAGACCTACCTTGATAGACGTGGTTTAAAGGTATGACCGAAGGTTGAACCGTCTC 1110  
Db 989 ACAGGGAGACTTATCTTACCTGCTCGCTATGAAGATATGATCGAGAGCGGTGAACCATCAC 1048  
QY 1111 AGTTGCTGCTGTGACATATTTGTACGTACAGTCGACCCCTTGAAGAGCCACTATCG 1170  
Db 1049 AGTCTGTTCTGTGTATGTTGTTAGTACAGTGGACCCATTTGAAGAGCCTCCCTTG 1108  
QY 1171 TCACTGCCAACACTGTCTATCCATTTCTGCTTGTGTTGATTTATCCGCTGGACAAAGTCTCTT 1230  
Db 1109 TTACAGCAAAACACAGTTCTCTCGATTTCTGTGAGTACCCGCTAGATAAAGTACGCT 1168  
QY 1231 GCTATGATCTGATGACGAGCTTCAATGCTGCACTTTTGACGCATTTGGCTGAGACTTCAG 1290  
Db 1169 GTTATGTTTCAGATGATGTTTCAGCTATGCTTACCTTTTGAATTCCTTTCTGAAACCGCTG 1228  
QY 1291 AGTTTCTAGGAATTTGGTACCATTTGTGAAGAAGTATGACATTTCAACCCAGAGCTCCCG 1350  
Db 1229 AGTTTCAAAAGAAATGGGTACCATTTTGCAGAAATTTCAACATTTGAACCTAGGSCCCTG 1288  
QY 1351 AGTTTACTTTTGGCAGAAATTTGATTTACCTTGAAGACAAAGTCCAGGCTTCATTTGTTA 1410  
Db 1289 AATCTTATTTTCCGCAAGATAGATTTACTTCAAGGACAAAGTCCAAACCGTCTTTTGTTA 1348  
QY 1411 AAGACCGCGCGCCATGAAGAGAGAAATATGAAGAAATTTAAATCAGGATAAATGCCCTAG 1470  
Db 1349 AAGAGCGAGCTATGAAGAGAGATATGAAGATTTAAAGTGAGGATAAATGCTCTG 1408  
QY 1471 TTTCTAAGCANTGAAGTCCCGAGGAGGATGGATCATCGAAGTGGCACACCATTCGC 1530  
Db 1409 TTGCCAAAGCAGAAATTCCTGGAAGAGGCTGGACAAATCGAGGATGGTACTCCCTGGC 1468  
QY 1531 CAGGAACAATACCGAGGATCATCTCGGAATGATTCAGGTTTCTCTTGGTCACAGTGGTG 1590  
Db 1469 CTGGTAACAACACTAGAGATCATCTCGAATGATACAGTGTCTTAGGCCATAGTGGG 1528  
QY 1591 GCCTTGATCTAGGATGATGATCTCCCGTTAGTTAGTGTCTCGTGAAGAGCTC 1650  
Db 1529 GTCTGATACCGTGAAGATAGCTTGCCTAGACTCATCTATGTTCTCGTGAAGAGGCG 1588  
QY 1651 CTGGTTCCAGCACCAAGAGGCTGGTGCATGAATGCCCTTGTGCTCTCAGCTG 1710  
Db 1589 CTGGATTTCAACACCAAAAAGGCTGGAGCTATGAATGCATTCGCTGATCTGCTG 1648  
QY 1711 TCCTTACTAATGGAATACATGTTGATCTTGATTTGATGATCATTACATCAACACAGCA 1770  
Db 1649 TTCTTACCAATGGAGCATATCTTTTGAACGTGGATTTGATCATTTACTTTAATAACAGTA 1708  
QY 1771 AGGCTGTCGAGAAGCTATGCTTCTTAATGATTCGAAGCTAGGTCGCGCAAGCTCTGTT 1830  
Db 1709 AGGCTATTAAAGAACTATGTTTTCATGATGAGCCCGCTATTGGAAGAGTGTCTGCT 1768  
QY 1831 ATGTGCACTCCCAAGAGTTGATGGGATTTGATAGGAATGATCGATATGCAACAGGA 1890  
Db 1769 ATGTGCACTCCCAAGCTTTTGAACGTGGATTTGATGATGATGATGATGATGATGATGAT 1828  
QY 1891 ACACGTCTTTTTTATTAACCTTGAGGGGCTTGAAGGCTTTCAAGGACCAAGTATTATG 1950  
Db 1829 ATATAGCTTTTTCGATATTAACATGAAGGTTGGATGGTATCCAGGTCAGTATATG 1888  
QY 1951 TGGGAACGTGTTGTTTTCACAGACAGCTATCTATGGTTATGAGCC-----CCCA 2003  
Db 1889 TGGGTACTGGTTGTTGTTTAAATAGGAGGCTCTATATGGGTATGATCTGTTTGTACGG 1948  
QY 2004 ATTAAGGCAAGAACCCAGGTTTCTTGGCATCA--CTATGTTGGGGGCAAGAAAGCAA 2061

Db 1949 AAGAAGATTAGAACCAATAAT 2008  
QY 2062 GCAAGTCAAGAACAGCATGTGACAGTGTCTG 2121  
Db 2009 AAAGTAGCAAGAAAGTATRACTA 2068  
QY 2122 TTCCAGTATTAATCTCGAAG 2181  
Db 2069 CTCACCTTTTCAATATGAGGA 2122  
QY 2182 AGAAATCAGTTCTCATGTCTCA 2241  
Db 2123 AGAGGTCTATTTCTAATGTCCCA 2182  
QY 2242 TTGTTGCCCTCCACTCTGATGGA 2301  
Db 2183 TTATTGGGCAACCTTTCATGGA 2242  
QY 2302 TTTTCAAGAAAGCTATCCATGT 2361  
Db 2243 TTCTGAAGAGGCTATTCATGT 2302  
QY 2362 CTGAGATTGGTTGGATCTATGG 2421  
Db 2303 AAGAGATTGGTTGGATCTATGS 2362  
QY 2422 AGCAAGAGGCTGGCGTTTCAGT 2481  
Db 2363 ATGCCCGGGTTGGATATCGAT 2422  
QY 2482 CCCCCATCAATCTTTTCAGATCG 2541  
Db 2423 CACCAATCAATCTTTTCATCG 2482  
QY 2542 AAATCTTTTCAGCGCGCATTGG 2601  
Db 2483 AGATTCTTTTAGCAGACATTTG 2542  
QY 2602 TGGAGAGATTCTCTTACATCAAG 2661  
Db 2543 TGGAGAGGATCGTTTATATCAAG 2602  
QY 2662 TCTATTGTATATTGCTGCTATC 2721  
Db 2603 CGATTGTATTCTTCCCGCTTTT 2662  
QY 2722 GCAACTTGGCCAGTATCTGGTT 2781  
Db 2663 GCAACTACGCGAGTATTTGGTTC 2722  
QY 2782 TTGAGATGAGTGGAGTGGTGT 2841  
Db 2723 TGGAGCTGAGATGGAGCGGTGG 2782  
QY 2842 TCATTGGAGGTATCTCTGCACAT 2901  
Db 2783 TCATTGGTGGACATCCGCCCAT 2842  
QY 2902 CCGGTATCGACACCAACTTCAC 2961  
Db 2843 CTGGTATCGACACCAACTTCAC 2902  
QY 2962 CTGAGCTCTACATGTTCAAGTGG 3021  
Db 2903 CAGAATCTACATCTTCAATGG 2962  
QY 3022 TTAACATGTTGGTGTCTGTGCT 3081  
Db 2963 TGAACCTCATAGGCAATGTGCT 3022  
QY 3082 GGGGCCGCTCTTTGGGAAGCTC 3141  
Db 3023 GGGGTCCGCTTTTCGGGAAGCTC 3082

Qy	3142	TCTCAAGGGTCTTATGGCAGGCAAACCGCACACCGAGGATTGTTCATGCTCGGCTG	3201
Db	3083	TCTTGAAAGGCTGTGTGGGAAGACAACCGAACCCATCGTCAITGTCTGTGCTG	3142
Qy	3202	TCTCTCGCTCTCATCTCTCTCTCTGTTGGTGTGGTGTGATCATCATT	3250
Db	3143	TTCCTTCGCCTCCATCTCTCTGTTGCTTGGTCAGGATCAATCCCTT	3191
RESULT 13			
LOCUS	AX030948	Sequence 11 from Patent WO9800549.	linear PAT 20-SEP-2000
DEFINITION	AX030948		
ACCESSION	AX030948		
VERSION	AX030948.1	GI:10278351	
KEYWORDS			
SOURCE	thale cress.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
TITLE	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
JOURNAL	1 (bases 1 to 3673)		
	Williamson, R.E., Peng L., Arioli, A. and Betzner, A.S.		
	Manipulation of cellulose and/or beta -1,4-glucan		
	Patent: WO 9800349-A 11 08-JAN-1998;		
	WILLIAMSON RICHARD EDWARD (AU); PENG LIANGCAI (AU); ARIOLI		
	ANTONIO (AU); UNIV AUSTRALIAN (AU); BETZNER ANDREAS STEFAN (AU);		
	COMMW SCIEN INT RES ORG (AU)		
FEATURES	Location/Qualifiers		
source	1..3673		
	/organism="Arabidopsis thaliana"		
	/strain="COLUMBIA"		
	/isolate="RSWI MUTANT"		
	/db_xref="taxon:3702"		
CDS	71..3316		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CAC09628.1"		
	/db_xref="GI:10278352"		
	/translation="MEASAGLVAGSVRRNELVRIHESDGGTKPLKWNNGIQICQIGD EDVDYDENVFVACNECAPPCVEYEERKDGTCCPOCKTRERRHRGSFRVEGGDD DVGVLLENFTFAQANKARHORHEEYSSSRHESQIPLLTHGHTVSGERTPTDP NSKAIEAKMGCMFAIGKKCYQQPQDFGIDLHDYARNRIVLFLNDCDHFI QGVPVTGCFGNRAQGLDLPVLTEDELEPNIIKVSCGGRKKKGSKKNYKEKRK LNRSNAPILFMHEDIDEGVEYDERSLMSORSVEKRFQSPVFAAATFMEGGGIP PTTPNALFEATHVTSQYEDKTENGKLIGMIVYSVEDILDITGPKHARGWLSIYN PPRAPGSPAPILSLDLNOYLWALGSIELLSRHCPFIWGYHGRLRLERTAYINT IYVIPISLIATCLIPAFCLDPIPELKSINASIWFILFSIAIVGILSELRWSSG VISBEDWRNQFWIGGTSAHLEFAVFGLLKLADINTFTVKATDEGDHAELYI FKWTALLPTLLTVLAGVSVAVNSQSWGLFLGKLFALMWLAHWLAHYPLFK GLLRQRNTPTIVVSVLLASIFSLLRVINRPFVDANPNANNFXGGVFE"		
BASE COUNT	961 a	759 c	868 g
ORIGIN			
Query Match	37.0%;	Score 1341.8;	DB 6; Length 3673;
Best Local Similarity	66.2%;	Pred. No. 4.9e-254;	
Matches 2097;	Conservative 0;	Mismatches 992;	Indels 80; Gaps 8;
Qy	92	GAGGCACGGGGCGGGGACGTGTGCAGATCTCGCGCGACGCCGCTCGGCACCACCGTTGGA	151
Db	163	GAGAATATGAATGCCAGATATGTCAGATCTGTGGTGATCATGTTGGACTCGCTGAAC	222
Qy	152	CGCGACAGTCTTACCGGCTCGACGCTGCGCGCTTCCCCTGCTCCGCCCTCGCTACGA	211
Db	223	TGAGATGCTTGTTCGGGTGAATGAATGTCCTTCCCTGTGTGTCGGCTTGTGATGA	282





QY 1651 CTGGTTCCAGACCAACAAGAGCTGGTGCATGAATGCCCTTGTTCGTGCTCAGCTG 1710  
DB 1659 CTGGATTCAACACCAACAAGAGCTGGAGCTATGAATGCATTGATCGTGTATCTGTTG 1718  
QY 1711 TCCTTACTATGACAAATACATGTTGAATCTTGATGTGATCAGTACATCAACAACAGA 1770  
DB 1719 TTCTTACCAATGAGGATATCTTTTGAACGTGGATTGTGATCATTAATCTTTATACAGTA 1778  
QY 1771 AGGCTGTCCGAAGAGCTATGTGCTTCTTAATAGATCCAAACCTAGTCCGCAAGTCTGTT 1830  
DB 1779 AGGCTATTAAGAAGCTATGTGTTTCATGATGACCCGGCTATTGGAAGAGTCTGCT 1838  
QY 1831 ATGTGAGTTCCACAAAGTTTGATGGATTGATAGGATGATCGATATGCAACACAGA 1890  
DB 1839 ATGTCCAGTTCCCTCAACGTTTGGAGTTATTTGATTTGACATCGATATGCCAACAGA 1898  
QY 1891 ACATCTGTTTTTGTATTAATTAACCTTGAGGGCTTTGAGGGCATTTCAAGGACCAAGTTATG 1950  
DB 1899 ATATAGTCTTTTCGATATTAACATGAAGGGTTGGATGGTATCCAGGGTCCAGTATATG 1958  
QY 1951 TGGGAAGTGTGTTCTTCAACAGAACAGCTATCTATGTTGTTATGAGCC-----CCCA 2003  
DB 1959 TGGGTAAGTGTGTTTAAATAGGACAGGCTATATGTTGGTATGATCCTGTTTGTAGCG 2018  
QY 2004 ATTAAGCGAAGAGCAGGTTTCTTGGCATCA--CTATGTGGGGCAAGAAGAGGCAA 2061  
DB 2019 AAGAAGATTTAGAACCAATAATATTGTCAGAGCTGTTGCGGGTCAAGGAAGAAAGTA 2078  
QY 2062 GCAAGTCAAGAAAGAGGCTCAGATAGAAAAGTCGACNAGCATGTGACAGTCTG 2121  
DB 2079 AAGTAGCAAGAAGTATACTACGAAAGAGGAGGATCAACAGAGTACCTCCCAAT 2138  
QY 2122 TTCCAGTATTCAATCTCAGAGACATAGAGAGGGTGTGAAGGTGCTGGGTTTGATGATG 2181  
DB 2139 CTCACATTTCAATATGGAGGACATCGATGAGGGTTTGAAGGT-----TATGATGATG 2192  
QY 2182 AGAATCAGTCTCATGCTCAATAGCCTTAGAAGAGATTGGCCAGTCAGCAGCAT 2241  
DB 2193 AGAGGTCTATTCTAATGTCCAGAGAGGTGTAGAAAGCGTTTGGTCAGTCCGCGGTAT 2252  
QY 2242 TTGTTGCCCTCCACTGATGGAATATGTGTGTTCCCTCAGTCCCTCCACTCCAGAAATCTC 2301  
DB 2253 TTATTCGGGCAACCTTCATGGAAGAAGCGGCATTCACCAACAACCAATCCCGCTACTC 2312  
QY 2302 TTATTGAAGAGCTATCCATGTCATAAGTTGTGGCTATGAGACAAAGTCTCAATGGGAA 2361  
DB 2313 TTCTGAAGAGGCTATTCATGTTAAGCTGTATGAGTGTGGTACGAAGAACTGAATGGGCA 2372  
QY 2362 CTGAGATTGGTTGGATCTATGGATCTGTACAGAAGATATCTTACTGGATTCAAGATGC 2421  
DB 2373 AAGAGATTGGTTGGATCTATGTTCCGTGACGGAAGATATCTTACTGGTTCAAGATGC 2432  
QY 2422 ACCAAGAGGCTGGGCTTCAGTCTATTCGATGCCCAAGCGCCAGCTTTCAAGGATCTG 2481  
DB 2433 ATCCCGGGGTTGGATATCGATCTACTGCAATCCCTCCAGCCCTGCGTTCAAGGGATCTG 2492  
QY 2482 CCCCCATCAATCTTTTCAGATCTGTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTTG 2541  
DB 2493 CACCAATCAATCTTTCTGATCTGTTGAACCAAGTCTTTCGATGGCTTTGGATCTATCG 2552  
QY 2542 AATTTCTTTTCAGCGGATTCGCCCTTATGTTATGGTATGGTACGAGGGGCGCTCAAGTTC 2601  
DB 2553 AGATTCTTTAGCAGACATCTCTATCTATCTGTTATGTTTACCATGGAAAGTTGAGACTTT 2612  
QY 2602 TGGAGAGATTGGTTACATCAACACCAACCATTTTACCCACTACCTCTCTCCGCTTCTAG 2661  
DB 2613 TGGAGAGGATCGTTATATCAACACCATGCTATCTCTTATACATCCATCCCTCTTATG 2672  
QY 2662 TCTATTGATATGCTGCTATCTGTCTGCTACATGGAAGTTTCATCATGCGCAGATTA 2721  
DB 2673 CGTATTGTATTCTCCCGCTTTTGTCTCTCATCACCAGCAGATTTTCATATACCCGAGATA 2732

QY 2722 GCAACTTGGCAGTATCTGGTT  
DB 2733 GCAACTACCGAGTATTGGTT  
QY 2782 TTGAGATGAGGTGGAGTGGTGT  
DB 2793 TGGAGCTGAGATGGAGCGGT  
QY 2842 TCATTGGAGGTATCTCTGCACA  
DB 2853 TCATTGGTGGACATCCGCCCA  
QY 2902 CCGGTATCGACACCAACTTCAC  
DB 2913 CTGGTATCGACACCAACTTCAC  
QY 2962 CTGAGCTCTACATGTTCAAGTG  
DB 2973 CAGAACTCTACATCTCAATG  
QY 3022 TTAACATGTTGGTGTGCTGCT  
DB 3033 TGAACCTCATAGCATATGTCCT  
QY 3082 GGGGGCGCTCTTTGGGAAGCT  
DB 3093 GGGTCCGCTTTTCGGGAAGCT  
QY 3142 TCTCAAGGTTCTTATGGCAGG  
DB 3153 TCTTGAAGGTTCTTGGGAAGG  
QY 3202 TCTCTCTGCTCTCTATCTTCTCC  
DB 3213 TTTCTTCTGCTCTCTATCTCTCG  
RESULT 15  
AF200532  
LOCUS  
DEFINITION  
AF200532  
ACCESSION  
AF200532.1  
VERSION  
AF200532.1  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
gene  
CDS

12 bp mRNA linear PLN 31-AUG-2000  
e-8 (Cesa-8) mRNA, complete cds.

Streptophyta; Embryophyta; Tracheophyta;  
ta; Liliopsida; Poales; Poaceae; PACC  
ogoneae; Zea.

dentjaris,T., Dhugga,K.S.,  
Delmer,D.P.,  
the plant cellulose synthase (Cesa) gene

13-1324 (2000)

S.,T.G.

onomic Traits, Pioneer Hi-Bred  
NW 62nd Avenue, Johnston, IA 50131, USA

1. .3812

/organism="Zea  
/db\_xref="tax  
/note="ident  
database cDNA  
Inc"

gene

CDS



```
/note="CESA-8"
/codon_start=1
/protein_id="AAF89968.1"
/db_xref="GI:9622888"
/translation="MEASGLVAGSHNRNELLVIRDRSGAAGGAARRABPQIC
GDEVGVGDGPFVACNECAFPVRCACYEYERREGSOAQPOCRTRYKRLKGGPRVAGD
BREDDVLEGEFLQDCAHEDDPOVAFSLRMOYSYGRGDHAGFPSPVNPVLL
TNGOMVDIPEQHALVPSYMSGGGGKKRIHPLPADNPLVOPSPMDSPKDLAAYG
YGSVAWKRMGKWKQRIQHVRSSEGGDWDGDDADLPLMDEARQPLSRKVPISSR
INPYRMILIVLVLGFFHYRVMHPAKDAFWLWLSVCEIWFAMSWILDQFPKWLP
IERETVILSLRFDKQPSQAPIDFFSYVDPTEKPLVNTANTVLSLSDYVEVE
KYSCVSDGAAMLTFEALSETFAKWKPFSEKNIEPRAPETFOQKIDYLDKDV
AASFVRRMRKREYERFKVINALVAKQVPEEGTWODGSPWPNVNRDHPQMIO
VFLSGSGRDVEGNEPLRLVYVRSKRPYHKKAGAMNALVRYSVAISNAAYLLNL
DGDHVINNSKAIKEMCFMDPLGVKVCYVQFQBEDGLDKNDYRANRVVFDIIM
KGLDGIQPIVYGVGCFVRQALGYDAPTKKPPSPSTNCWPKWLSLSCCSRNKK
KITPKTEKKRRLFKRAENPSAPALGEIDEGAPADIEKAGIVNQKLEKFGOSS
VFVASTLLENGTLKSPASPLSKAELHVSCEYEDKTDWGEIWIYGSITDILTG
FKMHCHWRSTYICPKRPAFGSAPLNSDLRHQVLRWALGSVEIFSKHCPWLYGY
GKLKLERFYSINSIVYPTWTSIPLLAYCTLPALCLLTGKFIPELTNVASIFWMAFI
CISVTGILEMWSGVAIDWNRNEQFWIGVSAHLFVAFQGLLKVFAGIDTSFTVTS
KAGDDEESELVTEKWTLLIPPTLLLFNFIQVAGISNAINNGYESWGLFGLFP
ARWTVHLVYPLKGLVGRNRTPIVIVWSILLASIFSLWVRDVPFLAKSNPLLEE
CGLDCN"
BASE COUNT      873 a   966 c   1111 g   862 t
ORIGIN
Query Match      37.0%; Score 1340.2; DB 8; Length 3812;
Best Local Similarity 65.5%; Pred. No. 1e-253;
Matches 2137; Conservative 0; Mismatches 1023; Indels 103; Gaps 8;
QY 89 CGGAGGACGGGGCGGGAGCGTGTGCAGATCTGCCCGCAGCGCGCTGGGACCCAGGTT 148
DB 309 CGCGGGCGCGGGCGGGAGCGCGTGTGCAGATCTGCCCGCAGCGCGCTGGGAGTT 368
QY 149 GGACGGCGACGCTTCCACCGCTGGGAGCTGTCCCGCTTCCCGGTCTCCCGCCCTGCTA 208
DB 369 CGACGGGAGCGCTTCTCGTGGCTGCAACGAGTGGCGCTTCCCGCTTCCCGCGCTGCTA 428
QY 209 CGACGACGAGCGGAGGAGGACCGGCGCTGTCCAGTGTCCAGTGTCCAGGACCAAGTACAGCG 268
DB 429 CGAGTACGAGCGCGGAGGAGGCTGCAAGCGTGTCCCGCGAGTGCAGGACCGCGCTACAGCG 488
QY 269 CCACGAGGAGCGGACGAGTCCGGGGGAGGAGGCGGACGACACTGATGCCCGATGATGG 328
DB 489 CCTCAAGGCGTCCCGCGGCTGGCGGCGAGGAGGAGGCGCGCTGCAGGACCTGGA 548
QY 329 TAGTGACTTCAAC---TACCCTGCATCTGGCACTGAGGACGAGGACGAGAGATGCTGA 385
DB 549 GGGCGAGTTGGCGCTCAGGAGCGGCGCGCCGACAGGAGGACGCGGAGTACGTCGCCGA 608
QY 386 CAGGATCGCAGCTGGCGCATGAACACCGGGGCGAGTGGCAATGTGGCCACCCCAAGTA 445
DB 609 GTCCATGCTCAGGCGCCAGATGAGCTACGGCGCGCGCGCGCGACGCGCACCCCGGCTTCAG 668
QY 446 TGACAGTGGGAGATCGGCTCTCCAAATATGACAGTGGAGAGATCCCTAGGGGATACGT 505
DB 669 CCCCGTCCCAACGCTCCCTCTCACCACACG-----GCCAGATGGTTGAT 714
QY 506 CCCTTTCAGTCACCAACACGAGATGTACAGGAGAAATCCCTGGAGCTTCGCTGATCATCA 565
DB 715 GACATCCCGCGGAGCAGCAGCGCTGCTGCGCTCTACATGAGCGGGCGGGCGGGG 774
QY 566 CATGATGTCCCTTACGGGGAACATCAGCAGAGCGTGTCCGTTTCCCTATGTGAATATTC 625
DB 775 GGCAAGAGGATCCACCCGCTCCCTTTCCGAGATCCCAACCTCCAGTGCACACCGAGATCC 834
QY 626 ACCAAATCCGTCAGG--GAGTTCCTCCGCGAGTATTTGGAATGTGCTCGAAGAGAGA 683
DB 835 ATGACCCGCTCAAGGATCTGCGCGCTACGGATATGCGACGCTGCGCTGGAAGGAGAGA 894
QY 684 GTTGATGGCTGGGAAATGAAGGAGACAAAGGCTGCATTCCTCATGACTAATATGGACAAGC 743
```

```
DB 895 ATGAGGGCTGGGAAGAGAGCAGGAGCGCTGCAGCATGTTCAGGAGCGGAGG----- 947
QY 744 ATTGCTCCCTCTGAAGGTGGGCGAGCTACTGACATCGATCGATCTACTGATPACAAATG 803
DB 948 -----TGGCGGTGATTTGGGATGGCAGCAT 972
QY 804 GAAGACGCTTTACTGAATGATGAAACTCGCCAGCCCTCTATCTAGAAAGTCCCCATCTCT 863
DB 973 GCAGATCTGCCCATTAATGGATGAAGCTAGGACGCCATTTCCAGAAAGTCCCTATATCA 1032
QY 864 TCCTCCAAAATAATCCCTACAGAAATGGTCTTCTCGCGTGTGTTGTTCTAAGCATC 923
DB 1033 TCAAGCGCAATTTAATCCCTACAGATGATATCGTTTATCCGTTGCTGGTTTGGGTTTC 1092
QY 924 TTCCTGCACATACGCTCTCACAAATCCTGTGCGTGAATGATACCCACTGTGGCTTTTATCT 983
DB 1093 TTCCTTCCACTACCGAGTGTGTCATCCGGCGAAGATGTCATTTGGCATTTGGCTCATATCT 1152
QY 984 GTTATATGTGAGATTTTGGTTTGTCTTATCCTTGGATCTGGATCAGTTCCTCCGAAGTGT 1043
DB 1153 GTAATCTGTGAAATCTCTGTTTGGATGTCTCTGGATTTCTTGTAGTCAAGTTCCTCAAGTGT 1212
QY 1044 CCAATCAACCGGAGAGCTTACCTTGTAGTACTGGCTTTTAAAGTATGACGGAAGGTGAA 1103
DB 1213 CCAATCGAGAGAGAGCTTACCTTGGACCGTTTGTCACTAAGGTTTGCAAGGAGGTCAA 1272
QY 1104 CCGTCTCAGTTGGCTGTGTGACATATTTGTTCAGTACAGTGCAGCCCTTTGAAGGAGCA 1163
DB 1273 CCCTCTCAGCTTGTCTCAATCGACTTCTTGTTCAGTACGTTTGTATCCCAAGAACT 1332
QY 1164 CCTATCTGACTGCCACACTGTGCTTATCCATCTTGTCTTGTGTTGATATCCCTGGACAG 1223
DB 1333 CCCTTGTGTCAGGGAACACTGCTTTCATCTTCTTGTGGATTTATTCAGGCTCGT 1392
QY 1224 GTCTCTTGTATCTGATGACGAGGATTCATCTGACTGCTGACTTTTGAAGCATTTGCTGAG 1283
DB 1393 GTCTCTGCTATGTTCTGATGATGCTGCTGCAATCTTACGTTTGAAGCATTTGCTGAA 1452
QY 1284 ACTTCAGATTTGCTAGGAATGGTACCATTTGTGGAAGATGATGACATCAAGCCAGA 1343
DB 1453 ACATCTGAATTTGCAAGAAATGGTTCCTTCCAGCAAAAGTTTAAATATCAGGCTCGT 1512
QY 1344 GCTCCGAGTCTTACTTTTCCAGAAAATTTGATTTACTGAAAGACAAAATCCAGGCTTCA 1403
DB 1513 GCTCTGAGTGGTACTTCCAAACAGAGATAGACTACTGAAAGACAAGGTTCTGCTTCA 1572
QY 1404 TTTGTTAAAGACCCCGGCGCATGAAGAGAAATATGAAGAATTTAAATCAGGATTAAT 1463
DB 1573 TTTGTTAGGAGAGGAGGCGGATGAAGAGAAATGAGTTCCCTGAGGAGTGAAGATGAAG 1632
QY 1464 GCCCTAGTTTCTAAGGCAATTTGAAAGTCCCGGAGGAAGTGGATCATCAAGATGGCACA 1523
DB 1633 GCCTTGTTGCAAAAGCCCAAAAGTTTCTTGAAGAGGATGGACAATGAAGATGAAGC 1692
QY 1524 CCATGGCCGCAACATACCGGATCATCTCGGAATGATTTTCAGGTTTCTTGGTCAAC 1583
DB 1693 CCCTGGGCTGGAACACGCTACGCGATCATCTCGGAATGATTCAGGATTTCTTGGCCAA 1752
QY 1584 AGTGGTGGCTTGTACTACTGAGGTTAATGAGCTCCCGCTTTAGTTTATGTCTCGTGA 1643
DB 1753 AGTGGGCTGCTGATGTGAAGGAAATGAGTTCCCTGCGCTGTTATCTCTCGAGAGA 1812
QY 1644 AAGGCTCTGGTTCAGCAGCAGAGGCTGGTGGCCATGATGATGCCCTTGTCTGCTGC 1703
DB 1813 AAGAGCCAGGTTATAACCATCAAGAGGCTGGTGGCCATGAATGCATGGTCTGCTGC 1872
QY 1704 TCAGTCTCTTACTACTGGAATACATGTTGAATCTTGTGATGATCAGTACATCAAC 1763
DB 1873 TCTGCTCTTATCAATGCTGCATACCTATTGAATCTGACTGTGATCAGTACATCAAC 1932
QY 1764 AACAGAGGCTGTCCGAGAGGATGTGCTTCTTAAATGATCCAAACCTAGTCCGCA 1823
```

Db 1933 AATAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGATCCTTTGGTGGGAAGAAA 1992  
Qy 1824 GTCTGTTATGTGAGTCCACAAAAGTTTGTATGGGATGATAGGAATGATGATATGCA 1883  
Db 1993 GTGTGCTATGTACAGTTCCTCAGAGTTTGTATGATGATGATGACAAAATGATGATACGCT 2052  
Qy 1884 AACAGGAACACTGTCTCTTTTGTATTAATCTTGAGGGCCCTTGACGGCATTCACAGACCA 1943  
Db 2053 AACAGGAACACTGTCTCTTTTGTATTAATCTTGAGGGCCCTTGACGGCATTCACAGACCA 2112  
Qy 1944 GTTTATGTGGAACACTGTCTCTTTTGTATTAATCTTGAGGGCCCTTGACGGCATTCACAGACCA 2003  
Db 2113 ATTTATGTGGTACTGTGATCTGTTTCACAGCGGCACTGTATGTTTATGATGCTGCT 2172  
Qy 2004 ATTAAGCGGAAGAACCA-----GGTTTCTTGGCAATCA 2036  
Db 2173 AAAACGAAGAAGCACCATCAAGAACTTGCAACTGTGCGCCCAAGTGGTGCCTCTCTTGC 2232  
Qy 2037 CTATGTGGGGCAAGAAAGCAAGTCAAAAGAAAGAGCTCAGATAAGAAAAA- 2095  
Db 2233 TGCTGCAGCAGGAACAAGATAAAGAAAGACTACAAAACCAAGACCGAGGAAGAGAAA 2292  
Qy 2096 -----GTCCGAACAGCATGTGGACAGTTCTGTTCAGTATTCATCTCGAAGACATAGAG 2150  
Db 2293 AGATTATTTTCAAGAAAGCAAAAACCATCTCTGCAATATGCTTTGGGTGAATTTGAT 2352  
Qy 2151 GAGGGTCTGAAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCTCATCTCAAAATGAGC 2210  
Db 2353 GAAGGTGCTCCAGGTG-----TGATATCGAAGAGCGCGGAATCGTAAATCAACAGAAA 2406  
Qy 2211 TTAGAAAGAGATTGGCCAGTCAGCAGCATTTGTTGGCTCCACCTCTGATGGAATATGTT 2270  
Db 2407 CTAGAGAAGAAATTTGGGAGTCTTCTGTTTGTGCGCATCAACACTTCTTGAGAACGGA 2466  
Qy 2271 GGTGTTCTCAGTCCCTCAGTCCAGAACTCTTTTGAAGAGACTATCCATGCTCATAGT 2330  
Db 2467 GGGNCCCTGAAGAGCCAGTCCAGTCTCTTCTGAGGAAGCTATACATGTTATCAG 2526  
Qy 2331 TGTGGCTATGAGACAAAGTCTGAATGGGAAGTGTGTTGGATCTATGGATCTGTC 2390  
Db 2527 TGGGGCTACGAAGCAAGACCCAGTCCAGTCTCTTCTGAGGAAGCTATACATGTTATCAG 2586  
Qy 2391 ACAGAGATATCTTACTGATTCAGATGCACGACAGAGGCTGGGCTTCAGTCTATATGC 2450  
Db 2587 ACAGAGATATCTTACTGATTCAGATGCACGACAGAGGCTGGGCTTCAGTCTATATGC 2646  
Qy 2451 ATGCCAAGGCCCAAGCTTCAAGGATCTGCCCACTCAATCTTTCAGATCGTCTGAAC 2510  
Db 2647 ATCCGAAGGCCCTGCATTCAAAAGTTCTGCGCCTCTGAACCTTTCGAGCCGCTTCAC 2706  
Qy 2511 CAAGTGTGGGGTGGCTCTCGGTTCTGTTGAAATTTCTTTCAGCGCATTCGCCCTTA 2570  
Db 2707 CAGGTCTGCTGGGCCCTTGGGTCGTCGAAATTTCTTCAGCAAGCACTGCCACTT 2766  
Qy 2571 TGTATGGCTACGGAGGGCCCTCAAGTTCCTGAGAGATTGCTTACATCAACACCAACC 2630  
Db 2767 TGTATGGGATACGGCGGGGCTAAAAATCTCTGGAAGGTTTCTTATATCAACTCCATC 2826  
Qy 2631 ATTTACCACTAACCTCTCTCCGGTCTCAGTCTATGATATGATTCCTGCTATCTGCTG 2690  
Db 2827 GTTTATCCCTGGACGTCCATCTCTCCGGCTTACTGTACCTTGCCCTGCCATCTGCCCTG 2886  
Qy 2691 CTCCTGGAAGTTCATGTCAGAGATAGCAACTTGGCCAGTATCTGGTTCATTGCG 2750  
Db 2887 CTCACGGGAAGTTTATCACACAGAGCTTACCAATGTCCGAGTATCTGTTTCAATGSCA 2946  
Qy 2751 CTCCTCTCTCAATTTCCCACTGATCTGATGAGTGGAGTGGTGGTGGTGGCATT 2810  
Db 2947 CTTTTCACTGCACTCCCGGACCGCATCTTGGAAATGAGGTGGAGTGGCGTGGCCATC 3006  
Qy 2811 GACAGTGGTGGAGGAATGAACAGTTCCTGGGTCAATGGAGGTATCTCTGCACATCTGTTT 2870  
Db 3007 GACGACTGGTGGAGGAACAGCAGTTCCTGGGTCTATCGGAGGGTTCGCGGCATCTGTTTC 3066

Qy 2871 GCGTCTTTTCAGGGTCTTCTGTA 2930  
Db 3067 GCGGTCTTCCAGGGCTGCTGA 3126  
Qy 2931 TCAAAGGCTAATGACGAAGAAG 2990  
Db 3127 TCAAAGCGGGGAGACAGCA-- 3183  
Qy 2991 CTCTCATCCCTCCGACGACCA 3050  
Db 3184 CTGCTGATACCCCGACACGCG 3243  
Qy 3051 TCCTAGCCCATCAACAGTGGTT 3110  
Db 3244 TCGAAGCGGATCAACAACGCGT 3303  
Qy 3111 GCCTTCTGGGTGATTTTTCAT 3170  
Db 3304 GCCTTCTGGGTGATCGTCCAC 3363  
Qy 3171 CGCACACGACGATTTGTCATCG 3230  
Db 3364 AGGACCGCGACGATCGTATCG 3423  
Qy 3231 TGGGTTCTGTTGATCCTATTC 3290  
Db 3424 TGGGTCGGCTGACCCGTTCC 3483  
Qy 3291 ATCAACTGTAGGAAGTGGGA 3  
Db 3484 CTGGACTGCAACTGAAGTGGG 6

Search completed: February 25, 2003 18:20:39  
Job time : 9088 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 14:53:29 : Search time 694 Seconds  
(without alignments)  
11766.204 Million cell updates/sec

Title: US-09-900-237-29

Perfect score: 3626

Sequence: 1 gcacgaggaaccccgctcca.....tcacatttggagaggtttt 3626

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2595	71.6	3799	24	AAS16458
2	2572.2	70.9	3746	21	AAZ99512
3	2572.2	70.9	3746	21	AAZ99527
4	2572.2	70.9	3773	21	AAZ99494
5	2325.2	64.1	3704	21	AAZ99533
6	1681.6	46.4	1734	21	AAZ58271
7	1677.6	46.3	3198	21	AAC49550
8	1673.6	46.2	3614	19	AAV06567
9	1408.4	38.8	3725	21	AAZ99500

10	1408.4	38.8	3725	21	AAZ99506	DNA encoding a mai
11	1408.4	38.8	3725	21	AAZ99521	DNA encoding a mai
12	1382	38.1	3851	21	AAA67114	Pinus radiata cell
13	1373	37.9	3753	21	AAZ99515	DNA encoding a mai
14	1373	37.9	3753	21	AAZ99530	DNA encoding a mai
15	1373	37.9	3780	21	AAZ99497	DNA encoding a mai
16	1343.4	37.0	3603	19	AAV06565	Arabidopsis cellu
17	1340.2	37.0	3786	21	AAZ58265	Corn cellulose syn
18	1340.2	37.0	3813	21	AAZ99509	DNA encoding a mai
19	1340.2	37.0	3813	21	AAZ99524	DNA encoding a mai
20	1338.6	36.9	3673	19	AAV06568	Arabidopsis cellu
21	1331.4	36.7	2830	24	AAZ16455	Corn cDNA encoding
22	1320.4	36.4	3776	21	AAZ58263	Corn cellulose syn
23	1317.2	36.3	3568	21	AAZ99491	DNA encoding a mai
24	1271	35.1	3936	21	AAZ58266	Corn cellulose syn
25	1271	35.1	3969	21	AAZ99503	DNA encoding a mai
26	1271	35.1	3969	21	AAZ99518	Soybean cellulose
27	1257	34.7	3517	21	AAZ58268	Cotton cellulose s
28	1217.2	33.6	3328	19	AAV34432	Cellulose synthase
29	1212.4	33.4	3207	19	AAV08372	Eucalyptus grandis
30	1201.8	33.1	3747	21	AAAG7145	Soybean cellulose
31	1183.4	32.6	2890	21	AAZ58270	Arabidopsis thalia
32	1177	32.5	3444	22	AAC83798	Cellulose synthase
33	1176.6	32.4	3828	19	AAV06566	Arabidopsis cellu
34	1147.2	31.6	3311	19	AAV08373	Cellulose synthase
35	1133.6	31.3	3232	22	AAC65448	Populus tremuloide
36	984.2	27.1	2248	19	AAV06562	Arabidopsis cellu
37	966.8	26.7	2125	21	AAZ58269	Soybean cellulose
38	933.6	25.7	2033	19	AAV08381	Cellulose synthase
39	921	25.4	2306	19	AAV34433	Cotton cellulose s
40	863.4	23.8	1401	21	AAZ44966	Arabidopsis thalia
41	835.4	23.0	2055	21	AAAG7111	Eucalyptus grandis
42	723.4	20.0	1860	21	AAAG7081	Eucalyptus grandis
43	666.2	18.4	1407	21	AAAG7086	Eucalyptus grandis
44	666.2	18.4	7234	22	AAC83247	Cellulose synthase
45	568.8	15.7	1741	19	AAV06569	Rice cellulose syn

## ALIGNMENTS

RESULT 1

AAS16458

ID AAS16458 standard; cDNA; 3799 BP.

XX AC AAS16458;

XX AC

DT 14-FEB-2002 (first entry)

DE Corn cDNA encoding cellulose synthase Cgrae19/cesa-19.

XX XX

KW Corn; ss: cellulose synthase; Cdpqs45; cesa-3; Cgrae19; cesa-9;

KW stalk quality; improved stand; silage; pericarp; kernel hardening;

KW handling ability; transgenic plant.

XX OS Zea mays.

XX Key Location/Qualifiers

XX CDS 238..3799

XX FT /\*tag= a

XX FT /product= "Cellulose synthase"

XX XX

XX WO200179516-A2.

XX XX

XX 25-OCT-2001.

XX XX

XX 12-APR-2001; 2001WO-US11951.

XX XX

XX 14-APR-2000; 2000US-0550483.

XX XX

XX (PION-) PIONEER HI-BRED INT INC.

XX XX

XX Dhugga KS, Helentjaris TG;









QY	1624	TAGTTTATGTCTCGTGAAAGCGTCTCTGGGTTCCAGCACCAAGAAAGGCTGGTGCCA	1683
DB	1879	TGGTCTATGTTTCTCGTGAAAGCGTCTGGATTCAGCATCATCAAGAAAGCTGGTGCCA	1938
QY	1684	TGAATGCCCTTGTTCGGTGTCACGCTGTCTTACTAATGGAACAATACATGTTGAATCTTG	1743
DB	1939	TGAATGCTCTTGTTCGGTGTCACGCTGTCTTACAAATGGAACAATACATGTTGAATCTTG	1998
QY	1744	ATTGTGATCACATACATCAACACAGCAAGGCTGCCAGAAAGTATGTGCTTCCTAATGG	1803
DB	1999	ATTGTGATCACTACATTAACAACAGTAAGGCTCTCAGGAAGCTATGTGCTTCCTATGG	2058
QY	1804	ATCCAAAGCTAGGTCGCCAAGTCTGTATGTGCAAGTTCACCAAGATTTGATGGGATTG	1863
DB	2059	ACCCTAACCTAGGAAGAGTGTCTGTACGTCCAGTTCCTCCAGAGATTGATGGCATTTG	2118
QY	1864	ATAGGAATGATCGATATGCAACACAGGAACACTGCTCTTTTGTATATTAACCTTCAGGGCC	1923
DB	2119	ACAGGAATGATCGATATGCCAACAGAAACCGTGTTCGATATATTAACCTTCAGAGTTC	2178
QY	1924	TTGACGGCATTCAGGACCAGTTTATGTGGAACTGGTTGTGTTTCAACAGAACAGCTA	1983
DB	2179	TTGATGGCATTCAGGACCAGTTTATGTGCGAACTGGCTGTCTTCAACCCGAACAGCTC	2238
QY	1984	TCTATGGTTATGAGCCCCCAATTAAAGCGAAGAACCCAGGTTTCTTGGCATCACTATGTG	2043
DB	2239	TATATGGTTTATGAGCCCCCAATTAAAGCAGAGAAGGTTGGTTTCTGTCTCACTATGTG	2298
QY	2044	GGGGCAAGAAGCAAGCAAGTCAAAGAAAGAGGCTCAGATAAGAAAGAACTCGCAACA	2103
DB	2299	CGGTAGGAAGAACAGCAAAATCAAGAA--GGGCTCGGACAGAAGACTCGCAGA	2355
QY	2104	AGCATGTGACAGTCTCTGTTCAGTATWTCAGAAAGCATAGAGAGGGTGTGGAAG	2163
DB	2356	AGCATGTGACAGTCTCTGTGCCAGTATTCAACCTTGAAGATATAGAGAGGGAGTTGAAG	2415
QY	2164	GTGCTGGGTTTCATGATGAGAAATCAGTCTCATGTCTCAATGAGCTTAGAGAAAGAT	2223
DB	2416	GCCTCGATTTGACGACGAGAAATCATCTTTATGTCTCAAGGAGCTCGAGAAAGAT	2475
QY	2224	TTGGCCAGTCAGCAGCATTTGTTCCTCCACTCTGATGGAAATATGGTGGTTTCCCTCAGT	2283
DB	2476	TTGGCCAGTCCGAGGTTTGTTCCTCCACTCTGATGGAGTATGTGGTGTTCCTCAGT	2535
QY	2284	CCTCCACTCCAGAACTCTTTGAAGAGCTATCCATGTCATAGTGTGGCTATGACGG	2343
DB	2536	CCGCACTCCGGAGTCTCTTCTGAAGAAGCTATCCATGTTATAGCTGTGGCTATGAGG	2595
QY	2344	ACAAGTCTGAATGGGAACTGAGATTGGTTGGATCTATGGATCTGTACACAGAAGATATC	2403
DB	2596	ACAAGACTGAATGGGAACTGAGATCGGTGGATCTACGGTCTCTGTGACAGAAGACATTC	2655
QY	2404	TTACTGGATTCAAGATGCACGCAAGGCTGGCTTCTAGTCTATATGATGCCCAAGGCC	2463
DB	2656	TCACCGGATTCAGATGTCAACGCGGAGGCTGGCGGTGATCTACTGTATGCCCAAGCGGC	2715
QY	2464	CAGCTTTCAAGGGATCTGCCGCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTGCGGT	2523
DB	2716	CAGCTTTCAAGGGTCTGCCGCCCATCAATCTTTTCGGACCGTCTCAACACAGTGTCTCGGT	2775
QY	2524	GGGCTCTCGGTTCTGTTGAATTTCTTTTCAGCGCGCATTTGCCCTTATGGTATGGCTACG	2583
DB	2776	GGGCTCTTGGGTTCCGTGGAGATCTCTTTCAGCGCGCATTTGCCCTTGTGTACGGCTACG	2835
QY	2584	GAGGGCGCTCAAGTTTCCTGGAGAGATTGCGTTACATCAACACACACATTTACCCACTAA	2643
DB	2836	GAGGGCGGCTCAAGTTTCCTGGAGAGATTGCGGTACATCAACACACACATCTACCCGCTCA	2895
QY	2644	CCCTCTCCGGCTTCTAGTCTATTGTATATTGCCCTGCTATCTGTCTGCTACCTGGAAAGT	2703
DB	2896	CGTCCATCCGGCTTCTCATCTACTGCATCTCTGCCCGGCATCTGTCTGCTACCGGAAAGT	2955

QY	2704	TCATCATGCCAGAGATTAGCAACTTGGCCAGATATCTGGTTCATTGGCGTCTTCTCTTCAA	2763
Db	2956	TCATCATTCAGAGATACGAACTTTCGCCAGCATCTGGTTCATCTCCCTCTTTCATCTCGA	3015
QY	2764	TTTTCGCCACTCGTATCCCTTGAGATGAGGTGGAGTGTGTGGCATTTGACAGTGGTGGGA	2823
Db	3016	TCCTGCCACAGGCGATCCCTGGAGATGAGTGGAGCGGGGTGGGCATCGACGATGGTGGGA	3075
QY	2824	GGAAATGAACAGATTCTGGGTTCATTGGAGGTATCTCTPGACATCTGTTTGGCGCTTTTCAGG	2883
Db	3076	GGAAACGAGCAGTTCTGGGTGATCGGGGCGATCTCGCGCACCTCTTGCGCGCTGTTCGAGG	3135
QY	2884	GTCTTCTGAAGGTGCTTGGCGGTATCGACACCAACTTCACCTGCACCTCAAAAGGCTAATG	2943
Db	3136	GCCTGTCTAAGTGTCTGGCGCGGCATCGACACCAACTTCACCGTCACCTCCAAAGGCTCTGG	3195
QY	2944	ACGAAGAAGCGCACTTTGCTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCCTC	3003
Db	3196	ACGAGNACGCGACITTCGCGGAGCTGTACATGTTCAAGTGGACGAGCGTCTCTGATCCCGC	3255
QY	3004	CGACGACCAATTTTGATCATTAACATGTTGGTGGTGGTCTGCTTGGCGACCTCTCACGCCATCA	3063
Db	3256	CCACCACCAATCTTGATCATCAACTGGTGGCGGTCTGCGCGGCATCTCTCACGCCATCA	3315
QY	3064	ACAGTGGTTACCAATCATCTGGGGCGCTCTTTGGGAAGCTCTTCTTGGCTCTCTGGGTGA	3123
Db	3316	ACAGCGGATACCAAGTCTGGGGCCCGCTCTTCGGCAAGCTCTTCTTCGCTCTTGGGTCA	3375
QY	3124	TTGTTCACTTATACCCATTCTCAAGGGTCTTATGGGCAGCAAAACCGCACACCGACGA	3183
Db	3376	TCGTCCACCTGTACCCGTTCTCTCAAGGGCTCATGGCAGGAGCAACCGCACCCCGACCA	3435
QY	3184	TTGTTCATGCTGGGCTGTCTCTCGCTTCTATCTTCTCTCTGCTGGGTTCGTGTTCG	3243
Db	3436	TCGTCTGCTGTGGGGCATCTCTGCTGGCGTCCATCTTCTCTTGTGTGGGTTCGCATTCG	3495
QY	3244	ATCCATTCACTACCCGCTCTCGCTGGCGCCAAATATCCAAACCTGTGGGCATCAACTGCTAGG	3303
Db	3496	ACCCCTTCAACACCCCGGTCACCTGGGCCCGGATACCCAGACGTGTGGCATCAACTGCTAGG	3555
QY	3304	AAAGTGGGAGTTT	3316
Db	3556	GAAGTGGGAAGTT	3568

RESULT 3  
AAZ99527  
ID AAZ99527 standard: DNA: 3746 BP.

XX  
PR 17-AUG-1998; 98US-0096822.

XX PA (PION-) PIONEER HI-BRED INT INC.

XX PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX  
DR WPI: 2000-224343/19.

DR P-PSDB; AAY84119.  
XX

PT New genes which encode maize cellulose synthase polypeptides in plants  
PT  
PT  
PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein -

PS Claim 1: page 176-181: 119pp: English: xx

CC The present sequence encodes a maize c

CC The cellulose synthase can be used for  
CC for improved stand or silage. It also

of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Sequence 3746 BP: 861 A: 976 C: 1006 G: 902 T: 1 other: XX

Query Match 70.9%; Score 2572.2; DB 21; Length 3746;  
Best Local Similarity 86.4%; Pred. No. 0;  
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2

QY	4	CGAGAACCCCGCTCCAGCTCTGTCTGCTGCGGTGGATCGCTCTGCGCGCCCATGG	63
Db	265	CCAGGCCCCAGGGTCCAGGCCAGCTCCCTCGACGTTTCTCGCGAGCTCGCTTGCCCATGG	324
QY	64	ACGGCCACGGGACGCCCTGAAGTCCGGAGGACGGGGCCGGGGACGTGTGCCAGATCT	123
Db	325	AGGGCACGCGGACGCGCTGAAGTCCGGGAGCGCGGTGGCGGACAGGTGTGCCAGATCT	384
QY	124	GCGCCACGGCTGGGCACACCGTTGGACGGCGGACGCTCTTACCAGCCCTGCGACGCTCTGCC	183
Db	385	CGCGCACGCGGTGGGCACACCGCGGAGGGGACGCTTTCGCGCCCTGCGACGCTCTGCG	444
QY	184	GCTTCCGGTGTGCGGCCCTGCTACGAGCAGGCGGAGGGGACCCAGGCGCTGCC	243
Db	445	GGTTCGGGTGTGCGGCCCTGCTACGAGTACGAGCGAAGGGACGGCAGCAGGCGTGCC	504
QY	244	TCCAGTGCRAAGACCAAGTACAAGCCGACAGAGGAGGCCACCGATCCGCGGGGAGGAG	303
Db	505	CCAGTGTCAAGACCAAGTACAAGCCGACAGAGGAGGCCCGCGCATCCGTGGGAGGAG	564
QY	304	GCAGCAGACTGATGCCGATGATGTAGTTCCTCAACTACCTTCGATCTGGCACTGAGG	363
Db	565	GAGACACACTGATGCCGA-----TAGCGACTTCAATTACCTTGCATCTGGCAATGAGG	618
QY	364	ACCAGAAGCAGAAGATTGCTGACAGGATGCGCAGCTGCGCATGACACCGGGGGCAGTG	423
Db	619	ACCAGAAGCAGAAGATTGCGGACAGAAATCCGAGCTGCGCATGAACGTTGGGGGACGC	678
QY	424	GCAATGTTGGCCACCCCAAGTAGACAGTGGCGAGATCGGCCTCTCCAAAGTATGACAGTG	483

Db 1759 GGATCATGCAAGATGGCACCATGGCCAGGAACAATACCMGGGACCATCCTCGAATGA 1818  
Qy 1564 TTCAAGTTTTCCTTGGTCCACAGTGGTGGCTTGATFACCTAGAGGTAATGAGCTCCCGGTT 1623  
Db 1819 TTCAAGTTTTCCTTGGTCCACAGTGGTGGCTTGATFACCTAGAGGTAATGAGCTACCCGTT 1878  
Qy 1624 TAGTTATGTGTCCTGCGTGAAGCGTCTCTGGTTCACACACACACAGAAAGCTGGTGCCA 1683  
Db 1879 TGGTCTATGTTTCTCGTGAAGCGTCTCTGGATTCACACATCACAGAAAGCTGGTGCCA 1938  
Qy 1684 TGAATGCCCTGCTGCTGCTCAGCTGCTTACTTAATGGACAATACATGTTGAATCTTG 1743  
Db 1939 TGAATGCTCTGTTGCTGCTCAGCTGCTTACCAATGGACAATACATGTTGAATCTTG 1998  
Qy 1744 ATTGTGATCACTACATCAACAACAGAGGCTGCTCCGAGAAGCTATGCTTCCCTAATGG 1803  
Db 1999 ATTGTGATCACTACATTAACAACAGAGGCTCTCAGGGAAGCTATGCTTCCCTAATGG 2058  
Qy 1804 ATCCAAACCTAGTCCGCAAGTCTGTTATGTCAGTTCCTCCACAAAGGTTTGATGGGATTG 1863  
Db 2059 ACCCTAACCTAGGAAGGAGTGTCTGCTAGCTCCAGTTCCTCCGAGAGATTCGATGGCAATTG 2118  
Qy 1864 ATAGGAATGATCGATATGCAACAGAGACACACTGCTTTTTCATATTAACCTTAGGGGCC 1923  
Db 2119 ACAGGAATGATCGATATGCAACAGAGACACCGCTGTTTTCGATATTAACTTGAGAGTCT 2178  
Qy 1924 TTGACGGCATTCAGGACCAAGTTTATGTGGGAACCTGGTGTGTTTTCACAGAACAGCTA 1983  
Db 2179 TTGATGGCATTCAGGACCAAGTTTATGTGGGAACCTGGTGTGTTTTCACAGAACAGCTC 2238  
Qy 1984 TCTATGTTTATGAGCCCCCAATTAAGGCGAAGAACCGAGGTTTCTTGCGATCACTATGCTG 2043  
Db 2239 TATATGTTTATGAGCCCCCAATTAAGCAGGAAGGAGGTTTCTTGTCATCACTATGCTG 2298  
Qy 2044 GGGCAAGAGGAGGCAAGCAAGTCAAGAAAGAGGAGCTCAGATAAGAAAGTCAACA 2103  
Db 2299 GCGGTAGGAAGAGGCAAGCAAAATCAAGAA -- GGGCTCGGACGAAGAAGTCCGAGA 2355  
Qy 2104 AGCATGTGACAGTTCCTGTCAGTATTCATCTCGAAGACATAGAGAGGTTTGAAG 2163  
Db 2356 AGCATGTGACAGTTCCTGTCAGTATTCATCTCGAAGACATAGAGAGGAGTTGAAG 2415  
Qy 2164 GTCTGGGTTGATGATGAGAAATCAGTCTCATGCTCAATGAGCTTAGAGAAGAT 2223  
Db 2416 GCGCTGGATTGACGACGAGAATACATCTTATGCTCAATGAGCTGGAGAGAGAT 2475  
Qy 2224 TTGGCAGTCAGCAGCAATTTGTCCTCCACTGATGGAATATGTTGGTGTTCCTCAGT 2283  
Db 2476 TTGGCCAGTCCGACGCTTTGTTGCTCCACTGATGAGATGTTGGTGTTCCTCAGT 2535  
Qy 2284 CPTCCACTCCAGATCTCTTTTGAAGAAGCTATCCATGTCATAAGTTGTGGCTATGAGG 2343  
Db 2536 CCGCACTCCGAGTCTCTTCGAAGAAGACTATCCATGTTATAGCTGTGGCTATGAGG 2595  
Qy 2344 ACAAGTCTGAATGGGAACCTGAGATTGGTTGGATCTATGGATCTGTACAGAGAATATTC 2403  
Db 2596 ACAAGACTGAATGGGAACCTGAGATCGGGTGGATCTACGGTTCTGTGACAGAAGACATTC 2655  
Qy 2404 TTACTGGATTCAGATGCAAGAGGCTGGGTTTCACTCTATTCATGTCATGCCCAAGCGCC 2463  
Db 2656 TCACCGGATTCAGATGCAAGCGGAGGCTGGCGGTGCGATCTACGATGCCCAAGCGGC 2715  
Qy 2464 CAGCTTTCAGGATCTGCCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTCCGGT 2523  
Db 2716 CAGCTTTCAGGATCTGCCCCCATCAATCTTTTCGACCGCTCTGAACCAAGTGTCCGGT 2775  
Qy 2524 GGGCTTCGGTTCCTGTTGAATTCCTTTTCAGCGGCAATTCGCCCTTATGGTATGGCTAGG 2583  
Db 2776 GGGCTTCGGTTCCTGTTGAATTCCTTTTCAGCGGCAATTCGCCCTTATGGTATGGCTAGG 2835  
Qy 2584 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCACTAA 2643  
Db 2836 GAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCACTCA 2895

Qy 2644 CCTCTCTCCGGCTTCTAGTCTATTTGATATATGCCTCTCTATCTGTCTGCTCACTGGAAAGT 2703  
Db 2896 CGTCCATCCGCTTCTCTCATCTACTGATCTGTGCCGCCATCTGTCTGTCTACCGGAAAGT 2955  
Qy 2704 TCATCATGCCAGAGATTAGCAACTTGGCAGATCTGTGGTTTCATTGCGCTCTTCTCTTCAA 2763  
Db 2956 TCATCATTCAGAGATCAGCAACTTCGCAGCATCTGGCAGCATCTGTCTCTCTCTCTCATCTCGA 3015  
Qy 2764 TTTTCCGCACTGGTATCTCTTGAGATGAGGTGGAGTGGTGTGGCATCTGTGGCTCTTTCAGG 2823  
Db 3016 TCTTCCGCACTGGTATCTCTTGAGATGAGGTGGAGTGGGCGGTGGGCATCTGAGTGGTGA 3075  
Qy 2824 GGAATGAACAGTCTGGTCACTTGGAGGATCTCTGCACATCTGTGTTGGCTCTTTCAGG 2883  
Db 3076 GGAACGAGCATGTTCTGGGTGATCGGGGCACTCTCCGCGACCTCTTTCGCGGTGTTCCAGG 3135  
Qy 2884 GTCTTCTGAAGTGTGTCGGGTATCGACACCAACTTCACTGTCACTCAAAAGGCTTAATG 2943  
Db 3136 GCTGCTCAAGGTGTGTCGGCGGATCGACACCAACTTCACTGTCACTCAAAAGGCTCTCG 3195  
Qy 2944 ACAGAAGCGCATCTTGTGAGCTCTACATGTTCAAGTGGAGCGCTTCTCATCTCCCTC 3003  
Db 3196 ACAGGACGGCATCTCGCGAGCTGTACATGTTCAAGTGGAGCGCTCTGTATCCCGC 3255  
Qy 3004 CGAGGACCATTTTCATCATTAACATGTTGGTGTGCTGTGGCAGCTCTCTACGCCATCA 3063  
Db 3256 CCAGCCACTCTGTGATCATCACTGTGCGGCGTGTGCGCGGCACTCTCTACGCCATCA 3315  
Qy 3064 ACAGTGTGTTACCAATATGCGGCGGCTCTTGGGAAGCTCTTCTTGTGCTTCTGGGTGA 3123  
Db 3316 ACAGCGATACCATGCTGTGGGCGGCTCTTCGCAAGCTCTTCTGCGCTTCTGGGTGA 3375  
Qy 3124 TTGTTCACTTATACCATCTCTCAAGGTTCTTATGGCAGGCAAAACCGCACACCCAGCA 3183  
Db 3376 TCGTCCACTGTACCGCTTCTCAAGGCGCTCATGGCAGGCAAGAACCCGACCCGACCA 3435  
Qy 3184 TTGTCATGCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3243  
Db 3436 TCGTGTGCTGTGGGCACTCTGCTGTGCGCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3495  
Qy 3244 ATCACTTCACTACCGCTCTCTGCGGCCAAATATCCAAACCTGTGGCATCACTGTGAGG 3303  
Db 3496 ACCCTTTCACACCGGCTCACTGCGCGGATACCCAGAGCTGTGGCATCACTGTGAGG 3555  
Qy 3304 AAGTGGGACTTT 3316  
Db 3556 GAAGTGAAGTT 3568

RESULT 4

AAZ99494  
ID AAZ99494 standard; DNA; 3773 BP.

XX AAZ99494;

AC AAZ99494;

XX 03-JUL-2000 (first entry)

DE DNA encoding a maize cellulose synthase.

XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker; ss.

XX Zea mays.

PH Key Location/Qualifiers

FT CDS 337..3565

FT /\*tag= a

FT /product= "cellulose synthase"

FT /transl\_except= (pos: 1817..1819, aa: Xaa)

FT /note= "no termination codon given; Xaa is an

FT unspecified amino acid"

XX









Qy	1418	CGGGCCATGAAGACAGAAATATGAAGAAATTTAAATCAGGATAAATGGCCCTAGTTCTCTAA	1477
Db	1621	CGGGCCATGAAGACAGAAATATGAAGAAATTTAAATCAGGATAAATGGCCCTAGTTCTCTAA	1680
Qy	1478	GGCATTGAAGTCCCGAGGAGGATGATCATCAAGATGGCACACCACTGGCGCAGGAAA	1537
Db	1681	GGCAAAAAGTTCGAGGAGGATGATCATCAAGATGGTACACCTTGGCTGGCGTGGGAA	1740
Qy	1538	CAATACAGGGATCATCTCGGAATGATTCAGGTTTTCTTGGTCACAGTGGTGGCCTTGA	1597
Db	1741	CAATACTAGGACCATCTCGGAATGATTCAGGTTTTCTTGGTCCACAGTGGAGGCTTGA	1800
Qy	1598	TACTAGGATTAATAGACTCCCCCGTTTATGTGTCGTCGTAAGAGCGTCTCGGGTT	1657
Db	1801	CGTTGAGCAATGAACCTTCCTCGTTTGTATGTGTCGTCGTAAGAGCGTCTCGGGTT	1860
Qy	1658	CCAGCACCAAGAGGCTGGCCATCAATCCCTTGGTCGTCGTCACAGTCTGCTCTAC	1717
Db	1861	CCAACATCACAGAGGCTGGCCATCAATGGCACTTGTTCGTCGTCATCAGTCTGCTCTAC	1920
Qy	1718	TAATGACCAATACATGTTGAATCTTGTATGTGATCACTACATCAACAACAGCAAGCTGT	1777
Db	1921	TAATGGCAATACATGTTGAATCTTGTATGTGATCACTACATCAATAATAGCAAGCTCT	1980
Qy	1778	CCGAGAAGCTATGTCTTCTTAATGGATCCAAACCTAGGTCGCAAGTCTGTGTTATGTGA	1837
Db	1981	TCGAGAAGCTATGTCTTCTTATGGACCAAACTAGGAAGAAATCTGTGTTATGTCCA	2040
Qy	1838	GTTCACCAAGGTTTGTAGGGATTTATAGGAATGATCGATATCAACAACAGCAACTGT	1897
Db	2041	ATTCTCCAGAGGTTTGTATGTTATGTATAGGAATGACCGATATGCAACAACAGCAACTGT	2100
Qy	1898	CTTTTTTGATTAATCTGAGGGGCTTGACGGCATTCAGGCCACTTATGTGGGAAC	1957
Db	2101	GTTTTTTCGATTAATCTGAGAGGCTTGACGGCATTCAGGGCCAGTTATGTGGGAAC	2160
Qy	1958	TGGTTGTGTTTTCAACAGAACACTATCTATGGTTTATGAGCCCCCAATTAAGCGGRAGA	2017
Db	2161	TGGTTGTGTTTTTAAACAGAACCGCTTATATGGTTTATGAGCTCCAGTCAAGAAAAAAA	2220
Qy	2018	GCCAGGTTCTTGCCATCACTATGTGGGGCAAGAAAGCAAGCAAGTCAAGAAAAAG	2077
Db	2221	GCCAGGTTCTTCTCTCTCGTTTGTGGGGCAAGAAAGCAAGTCAAAATCTAAGAA--	2277
Qy	2078	GAGCTCAGATAAGAAAAAGTCGAACAAGCATGTGACAGTTCTGTCCAGTATTCAATCT	2137
Db	2278	GAGCTCGGAAAAAGAAGTTCATAGACACGACAGACAGTTCTGTACCAGTATTAAATCT	2337
Qy	2138	CGAAGACATAGAGAGGTTGTGAAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCTCAT	2197
Db	2338	CGAAGATATAGAGGAAGGATTTGAAGGTTCTCAGTTTTGTATGAGAAATCGCTGATTAT	2397
Qy	2198	GTCTCAATCAGCTTAGAGAAGATTTGGCCAGTCAGCAGCATTTGTGCTCCACTCT	2257
Db	2398	GTCTCAATCAGCTTAGAGAAGATTTGGCCAGTCCAGTGTTTTGTAGGCTCTACTCT	2457
Qy	2258	GATGGAATATGGTGGTTCTCCTCAGTCTCCACTCCAGAAATCTCTTTTGAAGAAGCAT	2317
Db	2458	GATGGAATATGGTGGTTCTCCAAATCTGCAACTCCAGACTCTCTCTGAAAGAAGCAT	2517
Qy	2318	CCATGTCTAAGTTGTGGCTATGAGGACAAGTCTGAATGGGGAAGTGAAGTTGGTTGGAT	2377
Db	2518	TCATGTCTACGCTGTGGCTATCAGGACAAAACCTGACTGGGGAAGTGAAGTTGGGTTGGAT	2577
Qy	2378	CTATGGATCTGTACAGAAGATTTCTTACTGGATTCAGATTCACGCAAGAGCTGGCG	2437
Db	2578	CTATGGTTCTGTACAGAAGACATTTCTCACGGATTCAGATGCAATGCTCGAGGCTGGCG	2637
Qy	2438	TTCAGTCTATTGCAATGCCAAGCGCCAGTTTCAAGGGATCTGCCCCCATCAATCTTTC	2497
Db	2638	ATCAATCTACTGCAATGCTTAAGGACCAAGCTTTCAAGGGATCTGCTCTATCAACCTTTC	2697
Qy	2498	AGATCGTCTGAACCAAGTCTGCGGTGGGCTCTCGGTTCTGTGTGAATTTCTTTTCAGCGC	2557

[illegible]

FT XX /partial  
PN XX WO200004166-A2.  
XX XX 27-JAN-2000.  
PD XX  
XX XX 13-JUL-1999; 99WO-US15871.  
PF XX  
XX XX 14-JUL-1998; 98US-0092844.  
PR XX (DUPO ) DU PONT DE NEMOURS & CO E.I.  
XX XX  
XX XX  
PI Allen SM, Fader GM, Falco SC, Kinney AJ, Lightner JE, Miao G;  
PI Rafalski JA, Thorpe CJ;  
XX XX  
DR WPI: 2000-182431/16.  
XX P-PSDB; AAY58840.  
XX  
PT Nucleic acid fragments encoding cellulose biosynthetic enzyme useful as  
PT probes for isolating cDNAs and genes encoding homologous proteins, for  
PT producing transgenic plants -  
XX  
PS Claim 3; Page 66-67; 93pp; English.  
XX  
CC The present sequence is that of a contig of cDNA clones and encodes  
CC a portion (see AAY58840) of wheat cellulose synthase (CS). The cDNA  
CC clones were isolated from wheat 7-day seedling leaf and 7-day  
CC etiolated seedling root tissue cDNA libraries on the basis of  
CC homology to Arabidopsis and cotton CS sequences. The invention  
CC relates to isolated nucleic acid fragments encoding plant CS and to  
CC CS polypeptides. It also relates to the construction of a chimeric  
CC gene encoding all or a portion of the CS, in sense or antisense  
CC orientation, where expression of the gene results in altered levels  
CC of the CS in transformed host cells. The host cells can be used to  
CC screen compounds for their ability to inhibit CS activity. CS  
CC nucleic acids are also useful for producing transgenic plants having  
CC altered levels of CS, and hence altered levels of fibre. CS may  
CC also serve as a target for the development of novel herbicides.  
XX  
SQ Sequence 1734 BP; 441 A; 369 C; 420 G; 502 T; 2 other;

Query Match 46.4%; Score 1681.6; DB 21; Length 1734;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 1696; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1782 GAAGCTATGTCTCCCTAATGGATCCAAACCTAGGTCGCGCAAGTCTGTATGTGCAATTC 1841  
Db 2 GAAGCTATGTCTCCCTAATGGATCCAAACCTAGGTCGCGCAAGTCTGTATGTGCAATTC 61

Qy 1842 CCACAAAGGTTTGATGGATTGATAGGAATGATCGATATGCCAAGCAAGCACTGCTCTTT 1901  
Db 62 CCACAAAGGTTTGATGGATTGATAGGAATGATCGATATGCCAAGCAAGCACTGCTCTTT 121

Qy 1902 TTTGATATTAACCTTGAAGGCGCTTACGGCATTCAGGACCACTTATGTGGCACTGCT 1961  
Db 122 TTTGATATTAACCTTGAAGGCGCTTACGGCATTCAGGACCACTTATGTGGCACTGCT 181

Qy 1962 TGTGTTTCAACAGACAGATCTATCTATGTTATGAGCCCAATTAAGCGGAAGACCA 2021  
Db 182 TGTGTTTCAACAGACAGATCTATCTATGTTATGAGCCCAATTAAGCGGAAGACCA 241

Qy 2022 GGTTCCTTGGCATCATCTATGTTGGGCGAAGAAGCAAGCAAGTCAAGAAAGGAGC 2081  
Db 242 GGTTCCTTGGCATCATCTATGTTGGGCGAAGAAGCAAGCAAGTCAAGAAAGGAGC 301

Qy 2082 TCAGATAGAAAAGTGCACACAGCATGTGGACAGTTCTGTTCCAGTATTCATCTCGAA 2141  
Db 302 TCAGATAGAAAAGTGCACACAGCATGTGGACAGTTCTGTTCCAGTATTCATCTCGAA 361

Qy 2142 GACATAGAGGAGGTTGTTGAAGTGTGTTGATGATGAGAAATCAGTCTCATGTCT 2201  
Db 362 GACATAGAGGAGGTTGTTGAAGTGTGTTGATGATGAGAAATCAGTCTCATGTCT 421

Qy 2202 CAATGAGCTTAGAGAAAGAGAT 2261  
Db 422 CAATGAGCTTAGAGAAAGAGAT 481

Qy 2262 GAATATGTTGTTGTTCTCTCAGT 2321  
Db 482 GAATATGTTGTTGTTCTCAGT 541

Qy 2322 GTCATAAGTTGTTGCTATGAGG 2381  
Db 542 GTCATAAGTTGTTGCTATGAGG 601

Qy 2382 GGATCTGTACAGAAAGATATTC 2441  
Db 602 GGATCTGTACAGAAAGATATTC 661

Qy 2442 GTCATTATGTCATGCCCAAGCGCC 2501  
Db 662 ATCTATTGTCATGCCCAAGCGCC 721

Qy 2502 CGTCTGAACCAAGTGTGCGGTT 2561  
Db 722 CGTCTGAATCAAGTGTGCGGTT 781

Qy 2562 TGCCCTTATGTTATGTTGCTACG 2621  
Db 782 TGCCCTTATGTTATGTTGCTACG 841

Qy 2622 AACACACCAATTTACCACACTAA 2681  
Db 842 AACACACCAATTTACCACACTAA 901

Qy 2682 ATCTGTCTGCTCACTGGAAAGT 2741  
Db 902 ATCTGTCTGCTCACTGGAAAGT 961

Qy 2742 TTCATTGGGCTTCTTCTTTCAA 2801  
Db 962 TTCATTGGGCTTCTTCTTTCAA 1021

Qy 2802 GTTCGCTATGACGAGTGTGGAG 2861  
Db 1022 GTTCGCTATGACGAGTGTGGAG 1081

Qy 2862 CATCTGTTTGGCGTCTTTCAGG 2921  
Db 1082 CATCTGTTTGGCGTCTTTCAGG 1141

Qy 2922 ACTGTCACCTCAAGGCTAATG 2981  
Db 1142 ACTGTCACCTCAAGGCTAATG 1201

Qy 2982 TGGACGAGCTTCTCATCCCTCG 3041  
Db 1202 TGGACGAGCTTCTCATCCCTCG 1261

Qy 3042 GCTGGACCTCTACGCCCATCA 3101  
Db 1262 GCTGGTACCTCTACGCCCATCA 1321

Qy 3102 CTCCTTCTTGGCTTCTGGGTGAT 3161  
Db 1322 CTCCTTCTTGGCTTCTGGGTGAT 1381

Qy 3162 AGGCAAAACCGCACACCGAGAT 3221  
Db 1382 AGGCAAAACCGCACACCGAGAT 1441

Qy 3222 TCCTTGTGTGGGTTCTGTTGTA 3281  
Db 1442 TCCTTGTGTGGGTTCTGTTGTA 1501

Qy 3282 ACCTGTGGCATCAACTGCTAGG 3341

```

Db 1502 ACCGTGGCATCACTGCTAGGAAGTGGAGTTTGTAGAGACAGAAATATAACAGTGA 1561
QY 3342 TCGAGCAACAACCCGGGAGCCAGAGATATTTATGTGGGGTGTGAATTACTACGTTT 3401
Db 1562 TCGAGGACCACTGTGGAGCCAGAGATATTTATGTGGGGTGTGAATTACTACGTTT 1621
QY 3402 GAGAAGTTGTCAAATTCAGAAAAACACATTTGTAATAGATGTAATAGACCATCTACCG 3461
Db 1622 GAGAAGTTGTCAAATTCAGAAAAACACATTTGTAATAGATGTAATAGACTAICTACCG 1681
QY 3462 TTTTCATGAGTTAAGCTCTTCTTTTGGAAACAAAGGAA 3502
Db 1682 TTTTCATGAGTTAAGCTCTTCTTTTGGAAACAAAGGAA 1722

RESULT 7
AAC49550
ID AAC49550 standard; DNA; 3198 BP.
XX
AC AAC49550;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 61572.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
```

```
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 12-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 46.3%;
Best Local Similarity 72.0%;
Matches 2330; Conservative
QY 77 CGCCCTGAAGTCGCGGAGGCAC
DB 21 CGCGGAAAGCCGATGAAGAAC
QY 137 GGGCACACAGTTGGACGGGAC
DB 81 TGGCAAGACTGTTGATGGAGAT
QY 197 CGGCCCTGCTACGAGCAGCAG
DB 141 TCGGCTTTGCTACGAGTATGAG
QY 257 CAAGTACAAGGCCACACAGAGG
DB 201 CAGATACAAGAGGCTCAAGGT
QY 317 TGCCGATGATGTA---GTGAC
DB 261 AGCTGATGAAGTACTGTTGAG
QY 374 GAAGATTGCTGACAGGATGCC
DB 303 GAAATTTTCAGAGCGGATGCTT
QY 434 CCACCCCAAGTATCACAGTGGC
DB 363 GAAACCCCAAGTATGATAAGAG
QY 494 TAGGGATACGTCCCTTCAGTC
DB 414 GAGCAGACAAG-----
QY 554 GCCTGATCATCATGATGATGTC
DB 453 ACCTGAAGCCCTCTCTGTATCT
QY 613 --ATGTGAATCATTCACCAATG
DB 513 AGATGTCAATCAATCACCACAA
QY 671 CTGGAAGAGAGAGTTGATGGCT
DB 573 TTGGAAGGAGAGAGTTGATGGCT
QY 731 TAATGGGACAGCAATGCTCCCT
DB 633 CACGCAG-----GCTGCT
QY 791 TGAATACAACATGGAAGACGCT
DB 681 AGATATCCTAGCAGATGAGGCT
QY 851 AGTCCCATTTGCTTCTTCCCAAA
DB 741 AGTTTCAATTCCTTCATCACGG
QY 911 TGTTC TAAGCATCTTCTTGCACT
DB 801 TATCCTTTGCTCTTCTTTCAT
QY 971 GTGGCTTTTATCTCTTATATGTC
DB 861 ATGGCTGGTCTCTGTATATGTC
QY 1031 CCCGAAGTGGTTTCCAATCAACG
DB 921 TCCCAAGTGGTTTCTCTGTGAAC
QY 1091 CCGAGAAGGTGAACCGTCTCAGT
```

```
are 1677.6; DB 21; Length 3198;
ed. No. 0;
Mismatches 834; Indels 72; Gaps 8;
CGGGACAGTGTCAGATCTCGCGCACGGCCT 136
TCCGACAGACTGGCCAGATCTGTAGTGACATGT 80
CAGCGCTGGACAGCTGCGCGCTTCCGGGCTG 196
TGTGGCTTGTGATATTTGTTTCATTTCCAGTTG 140
GGAGGCAACCCAGGCTCTCCAGTGCAGAGAC 256
AGATGGGAATCAATCTTGTCTCAGTGCAGAAC 200
AGGATCCGGGGAGAGGAGGCGACAGACATGA 316
TGCTATTCTCGTGGTATTAAGACGAGGATGGCT 260
CTACCCTGCATCTGGCACTGAGGACCAAGCA 373
CTACCCT-----CAGAAGGA 302
GCCATGAACCCGCGGCGAGTGGCAATGTTGG 433
GCATCTTACTCGTGGGAAGGAGAGAAATGGG 362
CGGCTCTCCAAGTATGACAGTGGAGAGATCCC 493
-----TCTCTCAATCATCTTCTCGTCTCAC 413
CAGCCAGATGTCAGGAGAAATCCTTGAGGCTC 553
-----ATACTTCAGGAGAGTTTCTGCTGCTC 452
GGGAACATCAGCAGAGCTGCTCCGTTTCCCT - 612
TATCGCTGGGGAAGCGCTTCCCTATTATC 512
AAGGAGTCTTCCGGCAGTATTGGATGTTGC 670
GATTGGATCTCTGTTGGACTCGGGAATGAGC 572
AATGAAGCAGCAAGAGGTGCGATTCCCATGAC 730
AATGAAGCAAGAGAGAAATACTGCTCTGTCAG 632
AGTCCGGGAGCTACTGACATCGATGATCTAC 790
AAGAGTGGAG---TAGATATTGATGCCAGCAC 680
GAATGATGAACCTGCCAGCCTCTATCTAGAAA 850
GAATGAGAAAGCAGGAGGAGCCTCTGTCAGGAA 740
TCCTACAGAAATGTCATGTTCTGGGGTGGT 910
TCCTTACAGAAATGTTATTATGCTCGGGCTGT 800
CTCACAATTCCTGCGGTAATGATCAATCCACT 970
TATAACAACCCAGTCCCAATGCTTGTGCTCT 860
TTGGTTTGTCTTATCTCTGGATCTGATCAGTT 1030
TTGGTTTGCCTTATCTCTGGAATTTGGATCAGTT 920
SACTACCTTGTATGACTGGCTTTAAGGTATGA 1090
AACTACCTCGACAGGCTTGTCTTTAAGATATGA 980
TGCTGTTGACATATTTGTCAGTACAGTCGNCCC 1150
```







QY 1224 GTCTCTTGCTATGTAATCTGATGACGAGCTTCAATAGCTGACTTTTACGCAATTTGGCTGAG 1283  
D 1330 GTCTCTGTTATGTTCTGATGATGGTCTGCTATGTTATCATTTGAATCACTTGCAGAA 1389  
QY 1284 ACTTCAGAGTTTGTCTAGGAATGGCTACCAATTTGTGAAGAGATGATGACATTAACCCAGA 1343  
D 1390 ACATCAGAGTTTGTCTGTAATAGGTACCAATTTTGCAGAGAAATATACATAGAGCTCGT 1449  
QY 1344 GTCTCCGAGTTTACTTTTGGCAGAAATGATTACCTCGAAGACAAAGTCCAGCCTTCA 1403  
D 1450 GCACCAAGATGGTACTTTGCTCGAATAATAGATTACTTGAAGGATAAAGTTCCAGACATCA 1509  
QY 1404 TTTGTTAAAGACCGCGGCCCATGAAGAGAAATATGAAGAAATTAATAATCAGATAAAT 1463  
D 1510 TTTGTCAAAGATCGTAGAGCTATGAGAGGAAATATGAGGAATTTAAATATCCCAATCAAT 1569  
QY 1464 GCCTAGTTTCTAAGGCATTTGAAGTCCCGAGGAAGATGGATCATGCAAGATGGCACA 1523  
D 1570 GCATTTGTTTCCAAAGCCCTAAATATGCTGAAGAAAGGTTGGTTATGCAAGATGGCACA 1629  
QY 1524 CCATGCCAGGAAACAAATACCAAGGATCATCTGGAATGATTACAGGTTTTCCTTGTGTCA 1583  
D 1630 CCGTGGCTTGGAAATATACAGGGGACCATCCAGGAATGATCCAGGCTTCTTAGGGCAA 1689  
QY 1584 AGTGTGGCTGTGACTAGAGGTTAATGAGCTCCCGGTTTGTAGTTTATGTTGTCTGTGAA 1643  
D 1690 AATGGTGGACTTGTATGACAGAGGCAATGAGCTCCCGCTTGGTTATATGTTTCTCAGAA 1749  
QY 1644 AAGCTCTCTGGGTTCCAGCACCACAAAGAGCTGGTGCATGAATGCCCTTGTCTGTGTCT 1703  
D 1750 AAGGACCAAGGATCCAGCACCACAAAGGCTGGTCTATGATGACACTGGTGAGAGTT 1809  
QY 1704 TCAGCTGTCTTACTAATGGACAATACATGTTGAATCTTGATTTGATGATGACTACATCAAC 1763  
D 1810 TCAGCAGTTCTTACCAATGGACCTTTCATCTTGAATCTTGATTTGATGATCAATTAATAAT 1869  
QY 1764 AACAGCAAGCTGTCCGAGAAGCTATGTCTTCTCTAATGGATCCAAACCTAGTTCGCGCA 1823  
D 1870 AACAGCAAGCTTAAAGAGAACCAATGTCTCTGATGAGCACCACACCTCGGGAACAA 1929  
QY 1824 GTCTGTTATGTGAGTTCCTCCCAAAAGGTTTGAATGGATGATGAAGATGATGATGCA 1883  
D 1930 GTTGTGTTATGTTCAGTTCCCAAAAGATTTGATGATGATGATGATGATGATGATGATG 1989  
QY 1884 AACAGGAACAACGTCTTTTTCATATTAACCTTGAAGGGCTTGAAGGCATTAAGACCA 1943  
D 1990 AATCGTAATACCGTGTCTTTCATATTAACCTTGAAGGGCTTGAAGGCATTAAGACCA 2049  
QY 1944 GTTATGTGGGAACCTGTTTTCATGATGATGATGATGATGATGATGATGATGATGATG 2003  
D 2050 GTATATGTCCGAATGAGTGTGTTTTCATGATGATGATGATGATGATGATGATGATGATG 2109  
QY 2004 ATTAAGG- - - - -CGAAGAGCCAGGTTTCTTGGCATCACTATGTGGGGCAAGAAG 2057  
D 2110 ATAAAGTAAACACAGAAGCAAGTCTTTATCTAAGCTCTGTGGGATCAAGAAG 2169  
QY 2058 GCAAGCAAGTCAAGAAAGAGCTCAGATAGAAAAGTGCAGACAGCATGAGGACAGT 2117  
D 2170 AAGAATTCAAAGCTAAGAAAGAGTCCGACAAAAGAAATCAGGACGACATCTAGCTCA 2229  
QY 2118 TCTGTTCCAGTATTAATCTCAAGACATAGAGGAGGTTGTTGAAGTGTCTGGGTTTGTAT 2177  
D 2230 ACTGTTCTCTGATTAACCTCGATGACATAGAGGAGGTTGAAGTGTCTGGTTTGTAT 2289  
QY 2178 GATGAGAAATCAGTTCTCATGTTCTCAATGAGCTTAGAAGAGATTTGGCCAGTCA 2237  
D 2290 GATGAAAGGGCGTCTTAATGTGCGAAATGAGCCTTGGAGAGCGATTTGGACAGCTGTCT 2349  
QY 2238 GCATTTGTCCTCCACTCTGATGAATATGTTGGTGTCTCTCAGTCTCCACTCCAGAA 2297  
D 2350 GTTTTGTGCTCTACCTTAATGGAATGTTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2409

QY 2298 TCCTTTTGAAGAGCTATCCATGTCATAGTTGTGGCTATGAGGACAAGTCTGAATGG 2357  
D 2410 AACTTTCTCAAGAGGCTATCCATGTCATAGTTGTGGTATGAGGATGATGATGATGATG 2469  
QY 2358 GGAATCAGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2417  
D 2470 GGAATGAGATTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2529  
QY 2418 ATGACCAAGAGGCTGGGTTTCAGTCTATTCATGATGATGATGATGATGATGATGATGATG 2477  
D 2530 ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2589  
QY 2478 TCTGCCCTCAATCAATCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATG 2537  
D 2590 TCTGCTCTATCAATCTTTCAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2649  
QY 2538 GTTCAAAATCTTTCAGCGGCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2597  
D 2650 GTTGAAGATCTCTTCAGTGGCATGATGATGATGATGATGATGATGATGATGATGATGATG 2709  
QY 2598 TTTCTGAGAGATTCGCTTACATCAACACACCACTTACCCACTAACCTCTCTCCGCTT 2657  
D 2710 TTTCTGAGAGTTTGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2769  
QY 2658 CTAGTCTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2717  
D 2770 CTATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2829  
QY 2718 ATTAGCAATTTGGCCAGTATCTGCTTTCATGATGATGATGATGATGATGATGATGATGATG 2777  
D 2830 ATTAGTAACATTTGCAAGTATATGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2889  
QY 2778 ATCCTTGAAGATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2837  
D 2890 ATACTAGAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2949  
QY 2838 TGGTCTATTTGAGGATATCTCTGCACATCTGTTGGCTGTTTCCAGGGCTTCTCAGGGCTG 2897  
D 2950 TGGTCTATTTGAGGATATCTCTGCATCTGTTTATGCTGTTTCTCAAGGATCTCTCAAGATC 3009  
QY 2898 CTTGCCGATTCAGACCAACTTCACTGTCACCTCAAGGCTAATGACAGAGGCGAC 2957  
D 3010 CTTGCCGATTCAGACCAACTTCACTGTCACCTCAAGGCTAATGACAGAGGCGAC 3069  
QY 2958 TTTGCTGAGCTCTACATGTTTCAAGTGGAGAGCTTCTCTCCCTCCGACGACCATTTTG 3017  
D 3070 TTTGCTGAGCTCTACATGTTTCAAAATGGAACAACACTTCTCATTCGCCCAACGAGCTGCT 3129  
QY 3018 ATCATTAACATGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3077  
D 3130 ATTGTAACATTTAGTGGAGTGTGTCAGGAGTCTCTTATGCTATCAACAGTGGATACCA 3189  
QY 3078 TCATGGGGGCTCTTTTGGGAAGCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3137  
D 3190 TCATGGGGGCTCTTTTGGTAAGTGTGTTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3249  
QY 3138 CCAATTCCTCAAGGCTTATGGGAGGCAACACCGACGACGATCTCATCTGCTCTG 3197  
D 3250 CTTTCTCTCAAGGTTTGTGAGGTCAGACAGACCGGACCTTACCATTGTTGCTCTG 3309  
QY 3198 GCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3257  
D 3310 TCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3369  
QY 3258 GCTCTGCTGGCCCAATATCCAAACCTGTGGCATCACTGTAGGAA 3305  
D 3370 CGAGTCACTGGCCGGACATTTCTGGAATGTTGAATCACTGTTGAGAA 3417

RESULT 9  
AAZ99500  
ID AAZ99500 standard; DNA; 3725 BP.  
XX



QY 1250 AGCTTCAATGCTGACATTTTGGACGCATTTGGCTGAGACTTCAGAGTTTGTGCTAGGAATGGGT 1309  
D 1351 CTGAGTATGCTGACATTTTGGTCTCTCTGAAATGCGGAAATTTGCTAGAAAGTGGGT 1410  
QY 1310 ACCATTTTGAAGAAGTATGACATTTGAACCCAGAGCTCCGAGTTTACTTTTGGCCAGAA 1369  
D 1411 TCCTTTTGAAGAAGCAATATTTGAACCAAGAGCTCCAGAAATTTACTTTTGTCTCAAAA 1470  
QY 1370 AATTGATTACCTGAAAGACAAAGTCCAGCCTTCATTTGTTAAAGACCGCCGGCCATGAA 1429  
D 1471 AATAGATTACCTGAAGGACAAAATTCACCTTTCATTTGTTAAAGGAAAGACGACGAATGAA 1530  
QY 1430 GAGAGATATGAAGATTTAAATCAGGATAATGCCCTAGTTTCTAAGGCATTTGAAAGT 1489  
D 1531 GAGAGATATGAAGATTTCAAAATAGAATCAATGCCCTTGTGGCCAAAGCAGACAAAGT 1590  
QY 1490 CCCGAGGAGGATGATCATGCAAGATGGCACACCATGCCAGGAAACAAATACCAGGGA 1549  
D 1591 GCCTGAAGAGGGGTGGACCATGCTGATGGAATGCTTGGCCTGGGAATTAACCTTAGGGA 1650  
QY 1550 TCATCTGGAATGATTCAGGTTTTCCTTGGTCACAGTGTGGCCTTGATACGTAGGGTAA 1609  
D 1651 CCATCTGGCATGATTCAGGTTGTTTGGGGCACAGTGTGGCCTTGACACTGTGGA 1710  
QY 1610 TGAGCTCCCGCTTAGTTTATGTCCTGCGTGAAGAGCCTCCTGGGTTCCAGCACACAAA 1669  
D 1711 TGAATACCAAGTCTGTCTATGTCCTGCGTGAAGAGACCAAGCTTTCAGCATCANA 1770  
QY 1670 GAAGGCTGGTGCATGAATGCCCTTGTTCGTGTCACAGTGTGCCCTTACTAATGGACAATA 1729  
D 1771 GAAGGCTGGTGCATGAATGCCCTTGTTCGTGTCACAGTGTGCCCTTACTAATGGACAATA 1830  
QY 1730 CATGTTCAATCTGATTTGATCAGTACATCAACACAGCAAGGCTGCCGAGAGCAAT 1789  
D 1831 TCCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1890  
QY 1790 GTGCTTCTAATGATCCAAACCTAGGTCGCAAGTCTGTTATGTCAGTTCGCCCAAG 1849  
D 1891 GTGCTTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1950  
QY 1850 GTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1909  
D 1951 ATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2010  
QY 1910 TAATCTGAGGGCTTGAAGGATCAAGGACAGTTTATGTTGGGACAGTGTGTTGTTT 1969  
D 2011 CACATGAAAGGCTTAGTGCCATCAGGTCGAGTCTATGTTGGGACAGGATGCTGTTT 2070  
QY 1970 CAACAGAACAGTATCTATGTTTATGAGCC---CCCAATTAAGCGGAAGCCAGGTTT 2026  
D 2071 CAATAGCAGGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2130  
QY 2027 CTGAGCATCACTATGTTGGGGCAAGAAAGGCAAGCAAGTCAAAAGAAAGAGCTCAGA 2086  
D 2131 CATTTGTTTAAAGAGCTGCTGTTGTAAGAAAGAGAAAGAAAGAGTATATGATGATGATGAT 2190  
QY 2087 TAAGAAAAAGTCGAACAGCATGTTGACAGTCTCTGTTCCAGTATTCATCTCGAAGACAT 2146  
D 2191 TCAAGCCGTTATTAAGAGACAGAAATCTTCAGTCCCATCTTTAAGATGGAAGACAT 2250  
QY 2147 AGAGGAGGTTGTTGAAGGCTGCTGGGTTGATGATGAGAAATCAGTTCTCATGTCCTCAAT 2206  
D 2251 CGAGGAGGTTATGAAGGT-----TATGAGGATGAAAGTCAAGTCTTATGTCACAGAG 2304  
QY 2207 GAGCTTAGAAGAGATTTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTGTGATGAATA 2266  
D 2305 GAAATTTGGAGAAAGCTTTTGGTCAGTCTCCAAATCTTCATTTGCATCCACCTTTATGACTCA 2364  
QY 2267 TGTGCTGTTCCCTCAGTCTCCACTCCAGATCTCTTTTGAAGAAAGCTATCCATCTCAT 2326  
D 2365 AGGTGGCATCCAGCTTTCAACAAACCCAGCTTCTTACTGAAGGAAGCTATCCATGTTAT 2424  
QY 2327 AAGTTGTGGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGGTTGGATCTATGGATC 2386

D 2425 CAGCTGTGGGTACAGAGACAAAACCTGAATGGGAAAAGAGATTTGGCTGGATCTATGGTTC 2484  
QY 2387 TGTCAACAAGAGATTTCTTACTGGATTTCAAGATGCGAGGAGGCTGGGCTTCAGTCTA 2446  
D 2485 AGTTACAGAGATATTCTGACTGGTTTAAATGCAAGAGGCTGGCAATCAATCTA 2544  
QY 2447 TTGCATGCCCAAGCGGCCAGCTTTCAAGGATCTGCCCATCAATCTTCATGATGCT 2506  
D 2545 CTGCATGCCCAAGCGAGCTTTGTTCAAGGGTTCTGCACCAATCAATCTTCTGATGCT 2604  
QY 2507 GAACCAAGTGTGCGGGTGGCTCTCGTGTGAAATTTCTTTTTCAGCCGGCATTTGCC 2566  
D 2605 TAATCAGGTGCTCGTTGGGCTCTTGGTCAAGTGGAAATTTCTGCTAGCAGACATTTGCC 2664  
QY 2567 CTTATGATGATGGCTACGAGGAGGCGCTCAAGTTCTCTGGAGAGATTCGCTTACATCAAC 2626  
D 2665 TATATGTTATGGCTACAAATGGCGATTTGAAGCTTTTGGAGAGGCTGGCTTACATTAAC 2724  
QY 2627 CACCATTTACCCACTAACCTCTCTCCGCTCTAGTCTATTGATATTGCTGCTATCTG 2686  
D 2725 CATTTGTTATCCAAATCAATCTGTTCCGCTTATCGCCTATTTGTTGCTTCTGCTATCTG 2784  
QY 2687 TCTGCTCAGCTGGAAGTTTCATCATGCGAGAGATTTAGCAACTTTGCCAGTATCTGTTTCAT 2746  
D 2785 TCTTCTTACCAATAAATTTATCATTTCTGAGATTAGTAATATGCTGGAATGTTCTTCAT 2844  
QY 2747 TGGCCTCTCTCTTCAATTTTCGCCACTTGGTAICTTTCAGATGAGGTGGAGTGGTGG 2806  
D 2845 TCTTCTTTTGGCTCCCAATTTTCGCACTGTTATTTGGAGCTCAGATGGAGTGGTGG 2904  
QY 2807 CATTTGACAGTGTGGAGGAATGAACATTTCTGGGTCTATTGGAGGTATCTCTGCACATCT 2866  
D 2905 CATTTGAAGATTGTTGGAGAAATGAGCAGTTTGGGTTATTGGTGGCACCTTGCCTCATCT 2964  
QY 2867 GTTTGCCCTCTTTCAGGTCCTTTCAGAGGTTTTCGCGGTATTCGCGGTATTCGACACCACTTCAGTCT 2926  
D 2965 CTTCGCGTGTTCAGAGGCTCTGCTGAAAGTGTTCGCTGGGATGATACCAACTTCACAGT 3024  
QY 2927 CACCTCAAGGCTTAATGACGAAGAAGGCGACTTTGCTGAGCTCTACATGTTCAAGTGGAC 2986  
D 3025 TACCTCAAGGCTCTGATGAGGATGGCGACTTTGCTGAGCTATATGTGTTCAAGTGGAC 3084  
QY 2987 GACGCTTCTCATCTCCCTCCGAGGACCATTTTTCATTAACATGTTGTTGCTGCTGG 3046  
D 3085 CAGTTTGTCTATCCCTCCGACCACTGTTCTTGTTCATTAACCTGTCGGAATGGTGGCAGG 3144  
QY 3047 CACTCTCTACGCCATCAACAGTGGTTACCAATCATGCGGGCCGCTCTTTGGGAAGCTCTT 3106  
D 3145 AATTCGTATGCCATTAACAGCGGCTACCAATCTCGGGGTCCGCTCTTTGGAAGCTGTT 3204  
QY 3107 CTTTGCCTTCTGGGTGATTTGTTCACTTATACCCATTCCTCAAGGCTCTTATGGGACGGCA 3166  
D 3205 CTTCGATCTGGGTGATCTCCCATCTACCCCTTCTCAAGGCTCTCATGGGACGGCA 3264  
QY 3167 AAACGCGACACGAGATTTGTCATGCTGGGCTGTCCTCCTCGCTTCTATCTTCTCTT 3226  
D 3265 GAACGCGACGCAACAATCGTCATCGTTTGGTTCATCTCCTCTGCGTCTATCTTCTCTT 3324  
QY 3227 GCTGTGGGTTGCTGTTGATTCATTCACATPACC 3258  
D 3325 GCTGTGGGTGAAGATCGATCCTCTTTCATCTCCC 3356

RESULT 10  
AAZ99506  
ID AAZ99506 standard; DNA; 3725 BP.  
XX  
AC AAZ99506;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE DNA encoding a maize cellulose synthase.





Key	CDS	Location/Qualifiers
FH	FT	179..3397
FT	FT	/tag= a
FT	FT	/product= "cellulose synthase"
FT	FT	/note= "no termination codon given"
XX		
PN		WO200009706-A2.
XX		
PD		24-FEB-2000.
XX		
PF		16-AUG-1999; 99WO-US18760.
XX		
PR		17-AUG-1998; 98US-0096822.
XX		
PA		(PTON-) PIONEER HI-BRED INT INC.
XX		
PI		Dhugga KS, Helentjaris TG, Bowen BA, Wang X;
DR		P-PSDB; RAY84117.
XX		
FT		New genes which encode maize cellulose synthase polypeptides in plants useful for modulating the expression of cellulose synthase in plants and to produce transgenic plants expressing the novel protein -
XX		
PS		Claim 1; Page 160-165; 119pp; English.
XX		
CC		The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.
XX		
SQ		Sequence 3725 BP; 927 A; 807 C; 950 G; 1041 T; 0 other;
		Query Match 38.8%; Score 1408.4; DB 21; Length 3725;
		Best Local Similarity 66.8%; Pred. No. 0;
		Matches 2147; Conservative 0; Mismatches 966; Indels 99; Gaps 6;
QY		53 CCAGCGCATGACGGCAGCGGCCCTGAGTCCGGAGGCAGGGCCGGGAGGT 112 
Db		238 CCGCACACGCGGACGGCCGTGTCGGCGCTAAGCCCCCAAGAAGTGCGAATGGGCAGT 297 
QY		113 GTGCCAGATCTCGCCGACGGCCCTGGGACCACACTGTGGACGGCAGCTCTTCCACCGCCCTG 172 
Db		298 CTGCCAGATTGTGGCGACACTGTGGCGTTTCAGCCACTGGTGATGCTTTGTTGSCCTG 357 
QY		173 CGAGCTCTCGCGCTTCCGGCTCTCGCGCCCTGCTACGAGCAGCAGCGAAGAGGGGAC 232 
Db		358 CAATFAGTGTGCCCTGCCCTGCTCGCCGCCCTTGCTATGAGTACGACGCAAGGAAGGAA 417 
QY		233 CCAGGCTCGCTCCAGTGCACCAAGTACAAGCGCCACAGAGGGAGCCCGCATCCG 292 
Db		418 CCAATGTCGCCCTCAGTGCACAGACTAGATACAAGACAGAAGTAGCCCTCGAGTTCA 477 
QY		293 CGGGAGGAAGGGCAGCAGACTGATGCGCGATGATGGTGTAGTGACTTCAACTACCCCTGCATC 352 
Db		478 TGGTGATGATGAGGAGGAAGAT 533 
QY		353 TGGCACTGAGGACCAAGAACAG 412 
Db		538 CRAATGGGAAGGGCCAGAGTGC 574 
QY		413 CGGGGCGAGTGGCAATGTTGGG 472 
Db		575 GACGCTGATCTGTCTTTTCATCTG 634 
QY		473 GTATGACAGTGGAGAGATCCCT 532 
Db		635 AGTGGACA -----ACAGATATC 651 
QY		533 AGGAGAAATCCCTGGAGCTTCG 592 
Db		652 TGGAGAGATCCCTGATGTCATCG 711 
QY		593 CAGACGTGCTCCGTTCCCTAT 649 
Db		712 TGTGTATCCAAGGTTCCAGTT 771 
QY		650 CGGCACTATTGGGAATGTTGCC 709 
Db		772 CTATGGGCTTAATAGTGTGAC 831 
QY		710 CAAAGGTGCGATTCCCATGACT 769 
Db		832 CAAAATAATGTTGCAAGTACT 870 
QY		770 TACTGACATCGATGTCATCTACT 829 
Db		871 TAGAGGAGACATGGAGGGACT 930 
QY		830 TGGCCAGCCTCTATCTAGAAAT 889 
Db		931 AGCCCTACCTTTTGAGCCGCAAT 990 
QY		890 GGTCAATTGTTCTCGGTTGGT 949 
Db		991 AGTAATCAATCTCCGCTTTATC 1050 
QY		950 TGTGCTAATGCATACCCACTG 1009 
Db		1051 AGTGCCTAATGCTTATGGATTG 1110 
QY		1010 ATCTCGGATCTGGATCAGTTC 1069 
Db		1111 GTCTGGCTCTTAGATCAGTTC 1170 
QY		1070 TAGACTGGCTTTAAGTATGAC 1129 
Db		1171 CAGGCTTGCAATGAGTATGAT 1230 
QY		1130 ATTGTCACTACAGTCGACCCO 1189 
Db		1231 CTTTGTGCTACAGTGGATCCA 1290 
QY		1190 ATCCATTCTTGTGTTGATTAT 1249 
Db		1291 GTCCATTCTTGTGTGGATTAC 1350 



Db 1471 AATAGATTACCTGAAGGACAAAATTCACCTTCATTGTTTAAAGGAAAGACGAGCAATGAA 1530  
Qy 1430 GAGAGAAATATGAGAAATTTAAATCAGGATAAATGCGCTAGTTCTTAAGGCATTGAAAGT 1489  
Db 1531 GAGAGAGTATGAAGAAATCAAAATAGAATCAATGCGCTTGTGTGCCAAAGCAGAGAAAGT 1590  
Qy 1490 CCCCAGGAAAGGATGGATGCAAGATGCGACACCATGCCAGGAAACAAATACCAGGGA 1549  
Db 1591 GCCTGAAGAGGGTGGACCAATGGCTGATGGAATGCTTGGCCCTGGGAAATACCCCTAGGA 1650  
Qy 1550 TCATCCCTGGAATGATTCAGGTTTTCTTGGTGCACAGTGTGGCTTGTATCTAGAGGTTAA 1609  
Db 1651 CCATCCCTGGCATGATTCAGGTGTTCTTGGGGCACAGTGTGGCTTGACACTGATGAAA 1710  
Qy 1610 TGAGCTCCCCGTTTATGTGTCCTGTGAAAGCGTCCCTGGGTTCCAGCACACAA 1669  
Db 1711 TGAATTACCAAGCTTGTCTATGTCTCTCGTGAAGAGACCAAGGCTCTTACGATCACAA 1770  
Qy 1670 GAAGCGTGGTCCCATGAATGCCCTTGTTCGTGTCTCAGCTGTCTTACTAATAGGACAATA 1729  
Db 1771 GAAGCGTGGTGAATGAATGCATGATGATGCAATCAACAGCAAGCTGTCCGAGAAGCTAT 1789  
Qy 1730 CATGTTGAATCTTGATGTGATCACTACATCAACAGCAAGCAAGCTGTCCGAGAAGCTAT 1789  
Db 1831 TCTTCTCAATGTGATTTGTGACCAATTAATCAATAGCAGCAAGCTCTTAGAGAAGCAAT 1890  
Qy 1790 GTGCTTCCTAATGATCCAAACCTAGGTCCGCAAGTCTGTATGTGCAAGTCCACAAAG 1849  
Db 1891 GTGCTTCATGATGATCCAGCTCTAGGAAGGAAACCTTGTATGTACAAATTCACAAAG 1950  
Qy 1850 GTTTGATGGGATTCATAGGAATGATCGATATGCAACAGGAACACTCTCTTTTGTGATAT 1909  
Db 1951 ATTTGATGGCAATGACCTTGACGATCGATATGCTAATAGGACATAGCTCTCTTGATAT 2010  
Qy 1910 TAACCTGAGGGCCTTGACGGCATTCAGGACCAAGTATATGTGGGAACAGCTGTTGTTT 1969  
Db 2011 CAACATGAAGGCTTAGATGGCAATTCAGGTCCAGTCTATGTGGGAACAGGATGCTGTT 2070  
Qy 1970 CAACAGACACTATCATGTTATGAGCC---CCCAATTAAGCGGAAGACGAGCTTT 2026  
Db 2071 CAATAGGAGGCTTGTATGATGATATGCTGTTTGTGCTGAAGCTGATCTGGAACTAA 2130  
Qy 2027 CTGCGCATCACTATCTGGGGCAAGAAGGCAAGCAAGTCAAAAGAAAAGGAGCTCAGA 2086  
Db 2131 CATTTGTTCTTAAGAGCTGCTGTGTAGAGGAAGAGAAAGCAAGATTATATGGATAG 2190  
Qy 2087 TAAGAAAAGTCGAACAGCATGTGGACAGTTCGTGCCAGTATTCATCTCGAAGACAT 2146  
Db 2191 TCAAGCCGCTATTATGAAGAGAACAGATCTTCAGCTCCCATCTTTAACATGGAAGACAT 2250  
Qy 2147 AGAGGAGGCTGTGAAGGTGCTGGTGTGATGATGAGAAATCAGTTCCTCATGTCTCAAT 2206  
Db 2251 CGAGGAGGATTTGAAGT-----TATGAGGATGAAGGTCAGTGTATGTCCAGAG 2304  
Qy 2207 GAGCTTAGAGAGAGATTTGGCCAGTCAGCAGCATTTTGTGCCCTCCACTCTGTATGGGAATA 2266  
Db 2305 GAAATGGAGAAACGCTTTGGTCACTCTCCAATCTTCATTCATCCACCTTTATGACTCA 2364  
Qy 2267 TGGTGGTGTCTCAGTCTCCATCCAGAAATCTCTTTTGAAGAGCATATCCATGTCAT 2326  
Db 2365 AGGTGGCATACCACTCTCAAAACCCAGCTCTCTACTGAAGGAAGCTATCCATGTAT 2424  
Qy 2327 AAGTTGTGGCTATAGGACAACTGTAATGGGAAGTCTGAGATTTGGATCTATGGATC 2386  
Db 2425 CAGCTGTGGTACGAGGACAAACTGTAATGGGAAAGAGATTTGGCTGATCTATGGTTC 2484  
Qy 2387 TGTACAGAAGATATCTTACTGGATTCGAAGTGCACGCAAGAGCTGGCGTTTCAGTCTA 2446  
Db 2485 AGTTACAGAGGATATCTGACTGGGTTTAAATATGATCAAGAGGCTGGCAATCAATCTA 2544  
Qy 2447 TTGCATGCCCAAGCCCGCTTTCAGAGGATCTGCCCCCATCAATCTTTTCAGATCGTCT 2506  
Db 2545 CTGCAATGCCACACAGACCTGTTTTCAGAGGGTTCTGCAACCAATCAATCTTTCTGTCGTCT 2604

Qy 2507 GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTTGAAATTTCTTTTACGCCGCAATTGCCC 2566  
Db 2605 TAATCAGGTGCTCCGTTGGGCTCTTGGGTCACTGGAATTTCTGCTTAGCAGACATTTGCC 2664  
Qy 2567 CTTATGGTATGCTACGAGAGGCGCCTCAAGTTCCTGAGAGATTCGCTTACATCAACAC 2626  
Db 2665 TATATGGTATGCTACAATGGGCGATTTGAAGCTTTTGGAGAGGCTGGCTTACATTAACAC 2724  
Qy 2627 CACCATTACCCACATACTCTCTCCGCTCTAGCTATGCTATATGCTATGCTGCTATCTG 2686  
Db 2725 CATGTTTATCAATACATCTGTTCCGCTTATCGCTTATGCTGCTTCTGCTATCTG 2784  
Qy 2687 TCTGCTCACTGCAAAAGTTTCATCATGCCAGAGATTTAGCAACTTGGCCAGTATCTGTTTCA 2746  
Db 2785 TCTTCTTACCAATAAATTTATCATCTCGATAGTAAATATGCTGGAATGTTCTTCAT 2844  
Qy 2747 TCGGCTCTTCTTCAATTTTCGCCACTCGTATCTTGAATGAGATGAGGTGAGTGTGTTGG 2806  
Db 2845 TCTTCTTTTTCCTTTCGCACTGATATTTGGAGCTCAGATGGAGTGGTGTGTTGG 2904  
Qy 2807 CATTGACGAGTGTGGAGGAATGAACAGTCTGGGTTCATTGGAGGTATCTCTGCACATCT 2866  
Db 2905 CATTGAAGATTTGGTGAAGAAATGAGCAGTTTGGGTATTTGGTGGCACTCTGCCCATCT 2964  
Qy 2867 GTTTGCGCTCTTTCAGGGTCTTCTGAAAGGTGCTTGCCGATATCGACACCAACTTCACTGT 2926  
Db 2965 CTTGCGGGTGTCCAGGGTCTGCTGAAAGTGTGGCTGGGATTTGATACCAACTTCCACT 3024  
Qy 2927 CAGCTCAAGGCTTAATGACGAAGAGGCGACTTTTGTCTGAGCTCTACATGTTCAAGTGGAC 2986  
Db 3025 TACCTCAAGGCTCTCTGATGAGATGGGACTTTTGTGAGCTATATGTTTCAAGTGGAC 3084  
Qy 2987 GAGCTTCTCATCCCTCCGACGACCATTTTGATCATTAACATGTTGGTGTCTGTTGCTGG 3046  
Db 3085 CAGTTGCTCATCCCTCCGACCACTGTTCTGTCAATTAACCTGGTGGAAATGTTGGCAGG 3144  
Qy 3047 CAGCTCTAGCCATCAACAGTGGTTTACCAATCATCGGGGCGCCTTTTGGGAAGCTCTT 3106  
Db 3145 AATTCGTATGCCATTAAACAGGCGTACCAATCTCGGGTCCGCTCTTTGGAAAGCTGT 3204  
Qy 3107 CTTTGGCTCTGGGTGATTTGTCATCATACCCATTTCCCAAGGTCCTTATGGGACGCA 3166  
Db 3205 CTTCTCGATCTGGGTGATCTCCATCTACCCCTTCCCTCAAGGGTCTCATGGGACGCA 3264  
Qy 3167 AAACCCACACACGACGATGTCATCGTCTGGGCTGCTCCCTCGCTTCTATCTTCTCCTT 3226  
Db 3265 GAACCGACGCCCAACAAATCGTCATCGTTGGTCCATCTCCTTGCCTCTATCTTCTCCTT 3324  
Qy 3227 GCTGTGGGTTGCTGTTGATCCATTCATACCC 3258  
Db 3325 GCTGTGGGTGAAGATCGATCCTTTTCATCCTCC 3356

RESULT 12

AAA67114  
ID AAA67114 standard; DNA; 3851 BP.

XX  
AC AAA67114;

XX  
DT 31-OCT-2000 (first entry)

XX  
DE Pinus radiata cellulose synthase nucleotide sequence SEQ ID NO:69.

XX  
KW Eucalyptus grandis; pinus radiata; Monterey pine; modification;

XX  
KW plant cell wall; polysaccharide; polysaccharide biosynthetic pathway;

XX  
OS transgenic plant; ds.

XX  
PN Pinus radiata.

XX  
PN WO200022092-A2.

XX  
PD 20-APR-2000.

XX 08-OCT-1999; 99WO-NZ00169.  
XX 13-OCT-1998; 98US-0170862.  
XX 11-AUG-1999; 99US-0148426.  
XX (GENE-) GENESIS RES & DEV CORP LTD.  
XX (PLET-) FLETCHER CHALLENGE FORESTS LTD.  
XX Bloksberg LN;  
XX WPI; 2000-339328/29.  
XX P-PSDB; AAB16307.  
XX New genes encoding proteins involved in a plant polysaccharide  
XX biosynthetic pathway, useful for modulating or altering the  
XX polysaccharide content, composition or structure of the plant -  
XX Claim 1; Page 71-72; 301pp; English.  
XX The present invention describes isolated polynucleotides (PN) comprising  
XX a sequence selected from one of 835 nucleotide sequences given in  
XX AAA67073 to AAA67907, their (reverse) complements, sequences producing  
XX an Expectation (E) value of 0.01 or less compared to the 835 sequences,  
XX sequences at least 50% identical to them, 200, 100, 40 or 20-mers of the  
XX 835 sequences or sequences that are degenerately equivalent or allelic  
XX to the 835 sequences. The polynucleotides are used to modify the  
XX activity of a polypeptide involved in a polysaccharide biosynthetic  
XX pathway in the plant. They are especially used to modulate or alter the  
XX polysaccharide content, composition or structure of the plant. AAB16268  
XX to AAB16340 are proteins encoded by some of the polynucleotide sequence  
XX given in the present invention.  
XX Sequence 3851 BP; 1040 A; 796 C; 944 G; 1071 T; 0 other;  
Query Match 38.1%; Score 1382; DB 21; Length 3851;  
Best Local Similarity 66.4%; Pred. No. 0;  
Matches 2142; Conservative 0; Mismatches 1040; Indels 46; Gaps 9;  
QY 75 GAGCGCCCTGAAGTCCGGGAGGACGCGGGGCGGAGCTGTGCAGACTGTGCCCGCAGCGC 134  
DB 222 GAGGCGCGAAGCCTTTGACACGCTTGAGTGGCCAGCTGCGCAGATTTGTGGCGAGGAC 281  
QY 135 CTGGGACACACGTTGGACGCGGACGCTTTCACCGCCTGGGAGCTGTGCGGCTCCCGGTC 194  
DB 282 GTCCGGGCTTAACACAGACGCGGAGCTGTTCGCTGTGCTGTGAATCAGTGGGGTTTCTGTC 341  
QY 195 TGGCGCCCTGTACGACGACGAGCGCAAGGAGGACCCAGGCGCTGCTCCAGTGCAAG 254  
DB 342 TGTGCGCGCTGTGTGAGTACGAGACGAGAGGAAATCAGTGTGCGCGCAGTGCAAT 401  
QY 255 ACCAAGTACAGCGCCACAGAGGAGGAGCCAGCGATCCGCGGAGGAGGAGGAGGAGCAGCT 314  
DB 402 ACTCGTTACAAAGCGTCAAAAGGAGTGCA-----CGGGTGGAAGGTGACGATGAA 455  
QY 315 GATGCGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 374  
DB 456 GAAGACGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 515  
QY 375 AAGATTGCTGACAGGATGCGGAGCTGGCGCATGAACACCGGGGCGAGTGCGCAATGTTGGC 434  
DB 516 CAGATCACCGAGCGATGCTCCACGAGCAGATGAGCTATGCGGAGGTTCCCGACGACGAA 575  
QY 435 CAGCCCAAGTATG-ACAGTGGGAGATGCGGCTCTCCA-AGTATGACAGTGGAGATGCC 492  
DB 576 AATTGCGAGATGCTCATATATCAGAGCTTCTCCGCGAGATTCCTGCTGATCTTTCGAAACGGC 635  
QY 493 CTAGGGGATACGTCCTTCACTGACCAACAGCCAGATGTGAGGAGAAATCCCTGGAGCTT 552  
DB 636 CACTCGGTTGTAGTGGGAGATTCACAGCTCATCTACTAGCGAGACACCAATG----- 689  
QY 553 CGCCTGATCATCATGATGTGCTCCCTTACCGGGAACATCAGCAGAGCTGCTCGTTTCCT 612

DB 690 ---CTTGCCAAACCTGCAATGC  
QY 613 ATGTGAATCATTCACCAATTC  
DB 747 AGGATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT  
QY 673 GGAAGAGAGAGTGTGATGCTG  
DB 805 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG  
QY 733 ATGGGCAAGCATTTGCTCCCTC  
DB 865 CGGAAGGAGAGATATCAATATAA  
QY 793 AATACACATGGAAGAGGCTTT  
DB 925 ATCCGATATG-----C  
QY 853 TCCCATTTGCTTCTCCAAAT  
DB 976 TGCAATCTTCAAGCAAAAT  
QY 913 TTCTAAGCATTTCTTGCACATA  
DB 1036 TGTGGGTATTTTCTCCGCTA  
QY 973 GGCTTTTATCTGTATATGAG  
DB 1096 GGGCAGCTTCTATGCTTTGGA  
QY 1033 CGAAGTGGTTTCCAATCAACCG  
DB 1156 CCAAGTGGTTGCTATCAGTCG  
QY 1093 GAGAGGTGAACCGTCTCAGTTG  
DB 1216 GAGAAGCGAACCATCAATGCTT  
QY 1153 TGAAGGAGCCACCTATCGTCAC  
DB 1276 TGAAGGAGCCCTCTTGGTTAC  
QY 1213 CCGTGGACAGGCTCTCTTGCTAT  
DB 1336 CTGTAGACAATGTCTCTGTTA  
QY 1273 CATTTGGCTGAGACTTCAGAGTT  
DB 1396 CTCTCTGAGACTCAGAAAT  
QY 1333 TTGAACCCAGAGCTCCCGAGTTT  
DB 1456 TTGAGCCTCGCTCCCGAAATC  
QY 1393 TCCAGCCTTCATTTGTTAAAGAC  
DB 1516 TTCAACCCACTTTGTCAAHAG  
QY 1453 TCAGGATAAATGCCCTAGTTCT  
DB 1576 TGCATCAATCGTTGTTGCA  
QY 1513 AAGATGCCACCATCGCCAGGA  
DB 1636 AAGAGCTAGCGCTTGGCCTGGT  
QY 1573 TCTTGTGTACAGTGTGGCTTT  
DB 1696 TCTTGGTACAGTGGCGGCTC  
QY 1633 TGTCTGTGTAAGAGCTGCTGGG  
DB 1756 TTTCTGTGAGAGAGACCTGGT

QY 1693 TTGTTCTGCTCTCAGCTGCTCTTACTAATGAGCAATATACATGTTGAATCTTGAATGTTGATC 1752  
Db 1816 TGGTTTGGGGTTTCTGCTGTGCTACCAATGCTCCATTTTATCTGTAATCTGGAATGTTGATC 1875  
QY 1753 ACTACATCAACAACAGCAGGCTGTCGAGAGGCTATGCTGCTCCTAATGATGATCAAAACC 1812  
Db 1876 ACTACATTAACANTAGCAGAGGCAATCAGGAAGGCAATGCTTTATGATGATCCTCAGG 1935  
QY 1813 TAGGTCGCCAGCTCTGTTATGTCAGTTCCCAAAAGTTTGTATGGGATTCATAGGAATG 1872  
Db 1936 TTGGGAGAAAGCTGTTATGTCGAATCCCTCAGAGATTCGATGGTATTGATCGCAATG 1995  
QY 1873 ATCGATATGCAACAGGACACTGCTTTTGTGATTAATTAACCTGAGGGCCTTGACGCA 1932  
Db 1996 ACCGTTACGCCAATGAAACCGTATTCTTTGATATCAACATGAAGGCTGGAATGAA 2055  
QY 1933 TTCAAGGACCACTTATGTCGAACTGGTGTGTTTCAACAGAACAGCTATCTATGTTT 1992  
Db 2056 TTCAGGGCCTGTATATGTCGAACTGGATGCATGTTCAAGAGACAAGCTCTATATGGGT 2115  
QY 1993 ATGAGCCCCAATTAAGCGGAAGAGCCAGTTTCTTGGCATCACTATGTCGGGGCAAGA 2052  
Db 2116 ATGGGCTCCCAAGGCCCAAAACGCTCCCAAGATGGTGACCTGTGATTGT-----C 2166  
QY 2053 AGAAGCAAGCAAGTCAAGAAAGAGGCTCAGATAAGAAAGTCGAACAGCATGTGG 2112  
Db 2167 TCCCTTGTTCGGTCTCTGTAAGAGTCTCCGAAGAAATAGTAGCAGAAAGTGCAG 2226  
QY 2113 ACAGTTCTCTTCCAGTATTCAATCTCGAAGACATAGAGAGGGTGTGGAAGTGTGCGGT 2172  
Db 2227 GAATCCAGCTCCGCTTACAATCTGGACGGGATCGAGGAAGGATGAGAAGT-----T 2280  
QY 2173 TTGATCATGAGAAATCAGTTCTCAATGCTCAATGAGCTTACAGAGAGATTTGGCCAGT 2232  
Db 2281 ATGATGACGAAAGAGCATTTGTTATGAGCCAACTAGACTTCGAGAAGAGTTTGGCCAGT 2340  
QY 2233 CAGCAGATTGTTGCTCCACTCTGATGGAATATGGTGTGTTCTCAGTCTCCACTC 2292  
Db 2341 CTTACAGCTTTGTTCAATCCACTCTGATGGAGAAATGGTGTTCGCAACAGCAATC 2400  
QY 2293 CAGAATCTCTTTGAAGAAGCTATCCATGTCATAGTTGTGGCTATAGAGCAAGTCG 2352  
Db 2401 CAGCTGAATTTGTAAGGAGGCTATTCTATGTCATCAGCTGTGGATATGAAGCAAAACGG 2460  
QY 2353 AATGGGAAGTGAATGTTGGATCTATGGATCTCAGAGAGATATCTTACTGGAT 2412  
Db 2461 ANTGGGAAAGAGCTTGGATGGATCTATGGATCAGTCACAGAGGACATTCGACTGGAT 2520  
QY 2413 TCAAGATGCAAGCAGAGGCTGGCTTCAGTCTATTGCAATGCCAAGCGCCAGCTTTCA 2472  
Db 2521 TCAAGATGCACATCCAGGCTGGCGTCCATTACTGTATGCCAAGCGCAGCATTC 2580  
QY 2473 AGGATCTGCCCCATCAATCTTTTCAGATGCTGTAACCAAGTGTGCGGTGGGCTCTCG 2532  
Db 2581 AAGGGTCTGCTCAACAATCTATCAGACGGTTTGAACAGAGTGTGCGTGGGCTTTGG 2640  
QY 2533 GTTCTGTGAAATCTTTTCAGCGGCAATGCCCCCTTATGGTATGCTCAGGAGGCGCC 2592  
Db 2641 GATCAGTAGAAATTTTATGAGCAGACATTTGCCCAATCTGTTATGGCTATGGGGAGGTC 2700  
QY 2593 TCAAGTTCTTGGAGAGATTGGTTTACATCAACACCACTTATCCCACTAACTCTCTCC 2652  
Db 2701 TGAATGGCTTGAAGATTGGCTTATATCAACACCACTTGTCTATCCATTCACCTCTCTC 2760  
QY 2653 CGCTTCTAGTCTATGTATATGCTGCTATCTGCTCACTGGAAGTTTCATCATGC 2712  
Db 2761 CACTATTGCTTATGACACTTCCAGCCGCTCAGTTTGTCTCACTGGCAAAATTTGTGATCC 2820  
QY 2713 CAGAGATTAGCACTTGGCCAGTATCTGGTTCAATGGGCTCTTCCCTTCAATTTTCGCCA 2772  
Db 2821 CTCAGATCAGTACTTTTGCAAGTCTATTTTAAATAGTCTCTTTTCATCTCAATTTTGCCA 2880

QY 2773 CTGSTATCCTTGAGATGAGTGGAGTGGTGTGTCATTGACGAGTGGTGAGGAATGAAC 2832  
Db 2881 CTGSTATCTTGAAATGAGGTGAGTGGAGTGGATGAGCATTTGAGAAATGTCGGGAATGAAC 2940  
QY 2833 AGTTCTGGGTCATTGGAGGTATCTGTCACATCTGTTTGGCGTCTTTCAGGGTCTTCTGA 2892  
Db 2941 AGTTCTGGGTTATTGGAGGGTTCCTGCACATTTTTCGAGTTATTCAAGGCTGCTCA 3000  
QY 2893 AGTGTCTGGCGGTATCGACACCAACTTCACTGTCACCTCAAAGGCTTAATGAGCAAGAAG 2952  
Db 3001 AGTGTCTGGCAGGCAATGATACAAATTTACAGTCACTGCCAAGGC---ATCAGATGACG 3057  
QY 2953 GCGACTTTGCTGAGCTCTACATGTTCAAGTGGACGACGCTTCTCATCCCTCCGACGACCA 3012  
Db 3058 GTGAGTTTGGGAACTGTATGCAATCAATGAGCACACCTCTCATTCCTCTAGAACCC 3117  
QY 3013 TTTTGTGATCAATAACATGGTGGTGTGCTGTCGACACTCTACGCCATCAACAGTGGTT 3072  
Db 3118 TGCTTGTCAACAACCTTGTGGGGTGGTGTGGCGTAGCAGATGCAATCAACAATGGAT 3177  
QY 3073 ACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTCTTCTGCTTCTGGGTGATTTGTCAC 3132  
Db 3178 TTCAGTATGGGGTCCCTCTTGGGTAGCTTTTCTTTGTCATCTGGGTCTATTGTGACCC 3237  
QY 3133 TATACCCATTCCTCAAGGGTCTTATGGGAGGCAAAACCCACACACCGAGATTGTCAACG 3192  
Db 3238 TGTATCTCTTCTCAAGGGTCTCATGGCAGGCAAGAACCCACCATCTCGTGTGA 3297  
QY 3193 TCTGGGCTGCTCTCTGCTCTCATCTTCTTCTGCTGCTGGTGTGGTGTGATCAATCA 3252  
Db 3298 TTTGGTCAAAATCTCTGCTGCACTGTTTCTCTCTTTTCTGGGTAAAGAAATGATCCTTCT 3357  
QY 3253 CTACCGCTCTCGCTGGCCAAATATCAAACTGTGGCATCAACTGCT 3300  
Db 3358 TGAGTAAGTTAAAGGCCACATACTAAACAATGTGGCATCACTGCT 3405  
RESULT 13  
AAZ99515  
ID AAZ99515 standard; DNA; 3753 BP.  
XX AAZ99515;  
XX  
XX  
XX 03-JUL-2000 (first entry)  
XX DNA encoding a maize cellulose synthase.  
XX Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker; ss.  
XX Zea mays.  
FH Key Location/Qualifiers  
FT CDS 184..3405  
FT /\*tag= a  
FT /product= "cellulose synthase"  
FT /transl\_except= (pos: 916..918, aa: Xaa)  
FT /note= "no termination codon given; Xaa is an  
FT unspecified amino acid"  
XX WO200009706-A2.  
PN  
XX  
XX 24-FEB-2000.  
XX  
XX 16-AUG-1999; 99WO-US18760.  
XX  
XX 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
PI  
XX WPI; 2000-224343/19.



Db 1779 GAAGCGTGGTCAATGAATGGCGTGAATTCGTGTATCTGCTGTGCTGACAAATGGTGCCTA 1838  
Qy 1730 CATGTTGAATCTGATTTGATCAGTACATCAACAAGCAAGGCTGTCCGAGAAGCTAT 1789  
Db 1839 TCTTCTCAATGTGATTCGCGACCAATTAATCAATAGCAGCAAAAGCTTTAGAGAAGCAAT 1898  
Qy 1790 GTGCTTCTTAATGATCAAAACCTTAGTCCGCAAGTCTGTATGTGCAGTTCCCAAAAG 1849  
Db 1899 GTGCTTCATGATGATCGGCTCTAGGAAGAAACTGTGTATGTACAATTTCCACAGAG 1958  
Qy 1850 GTTTGATGGGATTGATAGGAATGATGATATGCAACAGCAAGCACTGTCTTTTGTATAT 1909  
Db 1959 ATTTGATGGCAATGACATTGCACGATGATGCTAATCGGAACATAGTTTCTTTGATAT 2018  
Qy 1910 TAACCTTGAGGGCCCTTCAACGCCATTCAGGACAGTTTATGTGGGACACTGGTTGTGTTT 1969  
Db 2019 CAACATGAAGGCTGTGATGGCAATTCAGGTCAGGTTTACGTGGGACAGGATGCTGTGT 2078  
Qy 1970 CAACAGAACAGCTATCTATGTGTTATGAGCC- - -CCCAATTAAGCGCAAGAACGAGGTTT 2026  
Db 2079 CAATAGACAGGCTTGTATGATACGATCCTGTTTCACTGAAGCTGATCTGGAGCCAA 2138  
Qy 2027 CTTGGCATCTATGTGGGCAAGAAGAGCAAGCAAGTCAAGAAAGAGAGCTCAGA 2086  
Db 2139 CATTTGTTATTAAAGAGCTGCTGTGTAAGGAAGAAAGAAAGCAAGATTATATGGATAG 2198  
Qy 2087 TAAGAAAAGTCGAACAGCATGGGACAGTCTGTGTCAGTATCAATCTCCAAGACAT 2146  
Db 2199 TCAAGCGGTATTATGAAGAGACAGAAATCTTCAGTCCCATCTCAATATGAAGACAT 2258  
Qy 2147 AGAGGAGGTTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAAT 2206  
Db 2259 CGAAGAGGTTATGAAGT- - - -TACGAGGATGAAGGTCAAGTCTGTTATGTCACAGAG 2312  
Qy 2207 GAGCTTAGAGAAGAGATTGGCCAGTCAGCAGACATTTGTTGGCTCCACTCTGATGAATA 2266  
Db 2313 GAAATTTGAGAAGAGCTTTGGTCACTCTCTATTTTTCGATCCACCTTTATGACACA 2372  
Qy 2267 TGGTGGTTCTCCTCAGTCCCTCACTCCAGAACTCTTTTGAAGAAGCTATCCATGTCAT 2326  
Db 2373 AGGTGGCATACCACTTCAACAAACCAGCTTCTCTACTAAGGAAGCTATCCATGTCAT 2432  
Qy 2327 AAGTTGTGGCTTAGGACCAAGTCTGAATGGGAACCTGAGATTGGTTGGATCTATGGATC 2386  
Db 2433 CAGTTGTGATATGAGCAAAACTGAATGGGAAAGAGATTGGCTGGATCTATGGTTC 2492  
Qy 2387 TGTACAGAAGATATCTTACTGATTCAGATGACGACGAAGGCTGGCGTTCAAGTCTA 2446  
Db 2493 AGTAACGGAGGATATCTGACTGGGTTTAAAAATGCATGCAAGGGGCTGGCAATCAATCTA 2552  
Qy 2447 TTGCATGCCAAGCGCCAGCTTTCAAGGATCTGCCCATCAATCTTTTCAGATCGTCT 2506  
Db 2553 CTGCATGCCACACAGCTTGTTCAGAGGTTCTGCAACCAATCAATCTTTCCGATCGTCT 2612  
Qy 2507 GAACCAAGTGTGCGGTGGGCTCTCGGTTCTGTGTTGAAATCTTTTTCAGCGGATTTGCC 2566  
Db 2613 TAATCAGTGTCTCGTGGGCTCTTGGCTCAGTGAATCTCTGTTAGTAGACATTTGCC 2672  
Qy 2567 CTTATGTTATGGCTACGAGGCGCCCTCAAGTTCCTCGAGAGATTCGCTTACATCAACAC 2626  
Db 2673 TATCTGTTATGGTTTACAAATGACGATTTGAAGCTTTTGGAGAGGCTGGCTTACATCAACAC 2732  
Qy 2627 CACCAATTTACCCTACCTCTCCCGCTCTAGTCTATGTTATATGCTGCTATCTG 2686  
Db 2733 TATTGTTATCCATACATCCATCCGCTTATGCTTATGTTGCTTCCCGCTATCTG 2792  
Qy 2687 TCTGCTCACTGGAAGTTCTATCATGCGCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCA 2746  
Db 2793 CTTCTTACCAATAAATTTATCATCTCTGAGATTAGCAATATGCTGGGATGTTCTTCAT 2852  
Qy 2747 TGGCGCTCTCTTTCAATTTTCGCCACTGGTATCTCTTGAAGATGAGGTGGGTGTGG 2806  
Db 2853 TCTTCTTTTCCCTCCATTTTGGCCACTGGTATATTTGGAGCTTAGATGGAGTGGTGTGG 2912

Qy 2807 CATTCAGCAGTGGTGGAGAAATGAAACATTTCTGGGTCAATTTGGAGGTATCTCTGCACATCT 2866  
Db 2913 CATTTGAAGATTTGGTGGAGAAATGAGCACTTTTGGGTTATTTGGTGGCACTCTGCCCATCT 2972  
Qy 2867 GTTTGCCGTCTTTTCAAGGTCTTCTGAAGGTCTTCCCGGTATCCGCGATCCGACACCAACTTCACGT 2926  
Db 2973 CTTCCAGTGTTCAGAGGCTCTGCTGAAAGTGTGGCTGGGATGATACCAACTTCACAGT 3032  
Qy 2927 CACCTCAAGGCTAATAGCAGAGAGGCGACTTTTGCTGAGCTCTACATGTTTCAAGTGGAC 2986  
Db 3033 TACCTCAAAAGGCATCTGATGAGGATGGGACTTTTCTGCTGAGCTATATGTTTCAAGTGGAC 3092  
Qy 2987 GAGCGTCTTCATCCCTCCGACGACCATTTTGATCATTAACATGTTGGTGTGCTGTCTGG 3046  
Db 3093 CAGTTGCTCATCTCTCCGACCACTGTTCTTGTCTAATTAACCTGGTCGGAATGGTGGCAGG 3152  
Qy 3047 CACCTCTACGCCATCAACAGTGGTTACCAATCATATGCGGCGGCTCTTTTGGAAAGCTCTT 3106  
Db 3153 AATTTCTTATGCAATTAACAGTGGCTACCAATCTCTGGGTCGCTCTTTTGGAAAGCTCTT 3212  
Qy 3107 CTTTGGCTTCTGGGTGATTTTCACTTATACCCATCTCTCAAGGTCTTATGGGCGAGCA 3166  
Db 3213 CTTCTCGATCTGGGTGATCTCTCAATCTCTACCCCTCTCTCAAGGTCTCTATGGGAAGCA 3272  
Qy 3167 AAACCGCACACCGACGATTTGTCATGCTCTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 3226  
Db 3273 GAACCGCACACCAACAACTGTCATTTGCTGGTCCATCTCTTTCATCTATCTCTCTCTCTCT 3332  
Qy 3227 GCTGTGGTTCGTGTGATPCCATTCACATACCC 3258  
Db 3333 GCTGTGGGTGAAGATCGATCCTTTTCATCTCCC 3364

RESULT 14  
AAZ99530  
ID AAZ99530 standard; DNA; 3753 BP.

XX AAZ99530;

DT 03-JUL-2000 (first entry)

XX DNA encoding a maize cellulose synthase.

DE Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker; ss.

OS Zea mays.

FH Key Location/Qualifiers

FT CDS 184..3405

FT /tag= a

FT /product= "cellulose synthase"

FT /transl\_except= (pos: 916..918, aa: Xaa)

FT /note= "no termination codon given; Xaa is an

FT unspecified amino acid"

XX WO200009706-A2.

PN 24-FEB-2000.

PD 16-AUG-1999; 99WO-US18760.

XX 17-AUG-1999; 98US-0096822.

PR (PION-) PIONEER HI-BRED INT INC.

PA Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

PI WPI; 2000-224343/19.

XX P-PSDB; AAY84120.

DR New genes which encode maize cellulose synthase polypeptides in plants  
XX

PT useful for modulating the expression of cellulose synthase in plants  
PT and to produce transgenic plants expressing the novel protein -  
XX  
PS Claim 1; Page 184-188; 119pp; English.

The present sequence encodes a maize cellulose synthase polypeptide. The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

SQ Sequence 3753 BP; 944 A; 805 C; 965 G; 1038 T; 1 other;

Query Match	37.9%	Score 1373;	DB 21;	Length 3753;
Best Local Similarity	66.3%;	Pred.	No. 0;	
Matches 2130; Conservative	0;	Mismatches	986;	Gaps 7;

53	QY	CCGGCCATGGACGGCAGCCGAGACGCCCTGAAGTCGCGGAGGACCGGGCCGGGGACGCT 112
243	Db	CCCCACAGACGGGGATGTGCGGGCTCGGGCTAAGCCCAACAAGAGTSCAAATGGACAGGT 302
113	QY	GTGCAGATCTGGCGCAGCGCCTGGGCACCAAGCTTTGGAGCGGCGAGCTCTTCACCGCCGTG 172
303	Db	CTGCCAGATTTGGCGTGACTCTCTGGGGTGTTCAGCCACTGGTGATGTCTTTGTTGCCCTG 362
173	QY	CGACGCTGTCCGCTTCCCCTGCTCCGCGCCCTCTCTACGAGCAGCGCAAGAGGAGGCAC 232
363	Db	CAATAGTGTGCCCTTCCCTGCTGCGGCCCATGCTATGATGATGACGCGCAAGGAGSGGAA 422
233	QY	CCAGGCGCTGCCTCCAGTGCNAAGACCAAGTATCAAGCCGCCACAGAGGAGGCCACGATCCG 292
423	Db	CCAAATGCTGCCCGCCAGTGCAGAGCTAGATACAAGACAGAAAGSTAGCCCTCGAGTTCA 482
293	QY	CGGGGAGGAAGGCGACACACTCATGCCGATGATGTAGTGTACTTCAACTACCCCTGCATC 352
483	Db	TGTTGATGAGGATGAGGAAGATGTTGATGACCTAGACAATGAATTCACATAC----- 534
353	QY	TGCGACTGAGGACCAAGAAGAGATTTGCTGACAGGATCGCGAGCTGGCGCATGAACAC 412
535	Db	-----RAGCAAGC 543
413	QY	CGGGGGAGTGGCAATGTTGGCCACCACCAAGTATGACATGGGGAGATCGGCTCTCCAA 472
544	Db	AGTGGCAAGGCCACAGTGGCAACTGCAAG---GAGATGATGCTGATCTGTCTTCATCT 600
473	QY	GTATGACAGTGGAGAGATCCCTAGGGATACGTCCTTCAGTACCACCAAGCCAGATGTC 532
601	Db	GCTCGCATGAGCCCATCATCGGATTCACGGCTGACAGGGGTCAA---CAGATATC 656
533	QY	AGGAGAAATCCCTGGAGCTTCGGCTGATCATCATGATGTCCCTTACCGGGGAACATCAG 592
657	Db	TGGAGAGATTCCTGATGCTTCCCTGACCGTCACTTATCCGAGTCCCAATCAGACTA 716
593	QY	CAGAGTGTGCTTTCCCTATGTGAATCATTCACCAAAATCCGTCAAGGAGTTC---TC 649
717	Db	TGTTGATCAAGCGTCCCAAGTTCCTGTGAGGATTTGGAGCCCTCCGAAGCACTTGAATTC 776



Db 1839 TCTTCTCAATGTCGATGGACCAATTACTTCAATACCAAGACTCTTAGAGAAAGCAAT 1898  
QY 1790 GTGCTTCTTAATGATCCAAACCTAGGTCCGCAAGTCTGTATGTGACGTTCCCAAAAG 1849  
Db 1899 GTGCTTCATGATGGATCCGGCTCTAGGAAGGAAACCTTGTATGTACAAATTTCCACAG 1958  
QY 1850 GTTGTGATGATGATAGGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1909  
Db 1959 ATTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2018  
QY 1910 TAACTTGAAGGGCTTGAAGGCAATCAAGGACAGTTTATGTGGGAACCTGTTGTGTTT 1969  
Db 2019 CAACATGAAGGCTGTGATGGCAATTCAGGTCAGTTTACGTGGGAACAGATGCTGTT 2078  
QY 1970 CAACAGAACAGTATCTATGTTTATGAGCC---CCCAATTAAGGGCAAGAACCCAGTTT 2026  
Db 2079 CAATAGACAGGCTTGTATGGATACGATCCTGTTTGAAGCTGATCTGGAGCCAAA 2138  
QY 2027 CTTGGCATCACTATGTGGGGCRAAGAGGCAAGCAAGTCAAGAAAAGAGCTCAGA 2086  
Db 2139 CATTTGTTATTAGAGCTGTGTGGTAGAAGGAAGAAAAGCAAGATTTATATGGATAG 2198  
QY 2087 TAAGAAAAGTCAAGCAAGCATGTGGACAGTCTCTGTCAGTATTCAAATCTCGAAGACAT 2146  
Db 2199 TCAAGCCGTATTATGAAGAGACAGATCTTCAGCTCCCATCTTCAATATGAAGACAT 2258  
QY 2147 AGAGGAGGCTGTGAAGGTCGTGGTTGATGATGAGAATCAAGTTCTCATGTCACAAAT 2206  
Db 2259 CGAAGAGGCTATTGAAGT-----TACGAGGATGAAGGTCAAGTCTTATGTCACAG 2312  
QY 2207 GAGCTTAGAAGAGATTTGGCCAGTCAGCAGCATTTGTCCTCCACTCTGATGGAATA 2266  
Db 2313 GAAATTCGAAGAAACGCTTTTGGTCAGTCTCCTATTTCATGTCATCCACCTTTATGACACA 2372  
QY 2267 TGGTGTGTTCTCCAGTCCCTCACTCCAGAACTCTTTTGAAGAGAGCTATCCATGTCAT 2326  
Db 2373 AGTGGCATACACCTTCACAAACCCAGCTCTCTACTAAGGAAGCTATCAAGTCAAT 2432  
QY 2327 AAGTTGTGGTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGTTGGATCTATGATC 2386  
Db 2433 CAGTTGTGGATATGAGGACAAACCTGAATGGGAAAGAGATTTGGCTGGATCTATGTTTC 2492  
QY 2387 TGTACAGAGATATCTTACTGATTCAGATGACGCAAGAGGCTGGGTTCACTCTA 2446  
Db 2493 AGTAACGGAGATATCTGACTGGGTTTAAATGCAATGCAAGGGGCTGGCAATCAATCA 2552  
QY 2447 TTGCATGCCCAAGCGCCAGCTTCAAGGATCTGCCCATCAATCTTTTCAGATCGTCT 2506  
Db 2553 CTGCATGCCACCCAGGACCTTGTTCAGGGTCTTCAGCAATCAATCTTCCGATCGTCT 2612  
QY 2507 GAACCAAGTGTGGGTGGGCTCTGGTCTGTGTTGAAATTTCTTTTCAGCGGGGATTCGCC 2566  
Db 2613 TAATCAGGTGCTCGTGGGCTCTGGGTCAAGTGAATTTCTGCTTAGTAGACATTTGCTC 2672  
QY 2567 CTTATGATGATGGCTACGAGGCGCTTCAAGTTCTCGGAGAGATTCGCTTACATCAACAC 2626  
Db 2673 TATCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2732  
QY 2627 CACCAATTAACCACTCACTCTCCCGCTTCTAGTCTATTTATATTTGCTGCTATCTG 2686  
Db 2733 TATGATATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 2792  
QY 2687 TCTGCTCACTGGAAGTTATCATCATGACGAGATAGCAACTTGGCCAGTATCTGGTTTCAAT 2746  
Db 2793 CTTCTTACCAATAAATTTATCATCTCTGAGATTAGCAATTTATGCTGGGATTTCTTCAT 2852  
QY 2747 TGCCTCTCTCTTCAATTTTCGCACTGCTATCTCTGATGATGATGATGATGATGATGATG 2806  
Db 2853 TCTTCTTTTCCCTCCATTTTGGCACTGATATATGAGCTTATGAGTGGAGTGGTGG 2912  
QY 2807 CATTGACAGTGGTGGAGGAATGAACAGTTCTGGGTCAATGGAGGTATCTCTGCACATCT 2866

Db 2913 CATTCAGATTTGGTGGAATGAGCAGTTTGGTATTGGTGGCACCTCTGCCATCT 2972  
QY 2867 GTTTGCCGTCTTTCCAGGCTCTTGAAGGCTCTTGAAGGCTTGGCGGTATCGACAACTTCACTGT 2926  
Db 2973 CTTGCCAGTGTTCAGGCTCTGCTGAAGTGTGGTGGATTTGATACCAACTTCAAGT 3032  
QY 2927 CACCTCAAGGCTAATGACGAAGAGGCGACTTTGCTGAGCTCTACATGTTTCAAGTGGAC 2986  
Db 3033 TACCTCAAGGCACTCTGATGAGATGGGACTTTGCTGAGCTATATGTTTCAAGTGGAC 3092  
QY 2987 GAGCGTTCTCATCTCCCTCGAGACCAATTTTGAATCAATTAACATGTTGTTGCTGCTGG 3046  
Db 3093 CAGTTGCTCATCTCTCGACCACTGTTCTTGTTCATTAACCTGTCGGAATGTTGCGCAG 3152  
QY 3047 CACCTCTACGCCATCAACAGTGGTTACCAATCATGAGGGGCCGCTCTTTGGGAAGCTCTT 3106  
Db 3153 AATTCTTATGCCAATTAACAGTGGCTACCAATCTGGGTCCTCTTTGGAAGCTGTT 3212  
QY 3107 CTTTGGCTTCTGGGTGATTTGTTCACTTATACCCATTTCTCAAGGCTCTTATGGGACGGCA 3166  
Db 3213 CTTCTGATCTGGGTGATCTCCATCTCTACCCCTTCTCAAGGCTCTCATGGGAAGGCA 3272  
QY 3167 AAACCGCACACCGAGATTTGTCATGCTGGGCTGTCCTCTCGCTTCTATCTTCTCTT 3226  
Db 3273 GAACCGCACACCAACAACTGTCATGCTGGTCCATCTCTTGCATCTATCTTCTCTCT 3332  
QY 3227 GCTGTGGGTTGCTGTTGATCCATTCACACCC 3258  
Db 3333 GCTGTGGTGAAGATCGATCTCTTTCATCTCTCC 3364

RESULT 15  
AAZ99497  
ID AAZ99497 standard; DNA; 3780 BP.  
XX  
AC AAZ99497;  
XX  
DT 03-JUL-2000 (first entry)  
XX  
DE DNA encoding a maize cellulose synthase.  
KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
KW transgenic plant; plant breeding marker; ss.  
XX Zea mays.  
FH Key Location/Qualifiers  
CDS 201..3422  
FT /\*tag= a  
FT /product= "cellulose synthase"  
FT /transl\_except= (pos: 933..935, aa: Xaa)  
FT /note= "no termination codon given; Xaa is an  
XX unspecified amino acid"  
PN WO200009706-A2.  
XX  
XX 24-FEB-2000.  
PF 16-AUG-1999; 99WO-US18760.  
XX  
PR 17-AUG-1998; 98US-0096822.  
XX  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
XX Dhugga KS, Helentjaris TG, Bowen BA, Wang X;  
XX WPI; 2000-224343/19.  
XX P-PSDB; AAY84109.  
XX  
XX New genes which encode maize cellulose synthase polypeptides in plants  
XX useful for modulating the expression of cellulose synthase in plants  
XX PT and to produce transgenic plants expressing the novel protein -  
XX

PS Claim 1; Page 97-102; 119pp; English.

XX The present sequence encodes a maize cellulose synthase polypeptide.  
CC The cellulose synthase can be used for the improvement of stalk quality  
CC for improved stand or silage. It also provides an increased concentration  
CC of cellulose in the pericarp, hardening the kernel and improving its  
CC handling ability. The sequences are used to produce transgenic plants  
CC and seeds expressing the cellulose synthase. The polynucleotide is  
CC used for modulating, preferably increasing, the level of the synthase  
CC in a plant cell. The plants are preferably monocots. The polynucleotide  
CC is also used as a probe or primer in the detection quantitation or  
CC isolation of gene transcripts. The probes are useful in detecting  
CC deficiencies in the level of mRNA in screenings for desired transgenic  
CC plant, for detecting mutations in the gene, for monitoring upregulation  
CC of expression or changes in enzyme activity in screening assays of  
CC compounds, for detection of any number of allelic variants of the gene,  
CC or for use as molecular markers in plant breeding programs. The  
CC isolated nucleic acids of the present invention can also be used for  
CC recombinant expression of their encoded polypeptides or for use as  
CC immunogens in the preparation and/or screening of antibodies. The  
CC proteins can be employed in assays for enzyme agonists or antagonists  
CC of enzyme function or for use of immunogens or antigens to obtain  
CC antibodies specifically immunoreactive with a protein.

XX  
SQ Sequence 3780 BP; 946 A; 817 C; 976 G; 1040 T; 1 other;

Query Match 37.9%; Score 1373; DB 21; Length 3780;

Best Local Similarity 66.3%; Pred. No. 0;

Matches 2130; Conservative 0; Mismatches 986; Indels 96; Gaps 7;

QY 53 CCGCGCCATGGACGGCGGACGCCCTGAAGTCCGGGAGGCACGGGGCGGGACGT 112  
DB 260 CCGCGCAGCAGCGGATGTCGGGGGTGCGCTAAGCCCAAGAGGTGCGAATGGACAGT 319  
QY 113 GTGCCAGATCTGCGCGGCGGCTGGGACACACGCTTGGAGCGGACGCTTTCACGCCCG 172  
DB 320 CTGCCAGATTTGGGTGACTCTGTGGGTGTTTCAGCCACCTGGTGATGCTCTTTGTCCTG 379  
QY 173 CGACGTCTCCCGCTTCCCGGCTGCGGCCCTGCTACGAGCAGCAGCGCAAGAGGGCAC 232  
DB 380 CAATGAGTGTGCTTCCCTGTGTCGCCCATGCTATGATGATGAGCGCAAGGAGGGAA 439  
QY 233 CCAGGCGCTCCCTCCAGTCAAGACCAAGTACAAAGCCGACAGAGGAGCCGACGATCG 292  
DB 440 CCAATGCTGCCCGCCAGTGAAGACTAGATACAAAGAGACAGAAAGGTAGCCCTCGAGTTCA 499  
QY 293 CCGGGAGGAAGCGACGACACTGATGCCGATGATGGTAGTGACTTCAACTACCTGTCATC 352  
DB 500 TGGTGATGAGGATGAGGAAGATGTTGATGACCTAGACATGAATCAACTAC----- 551  
QY 353 TGGCAGTGAAGGACCAAGACAGAGATGCTGACAGGATGCGCAGCTGGCGCATGAACAC 412  
DB 552 -----AAGCAAGGC 560  
QY 413 CCGGGCAGTGCATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCGCTCTCCAA 472  
DB 561 AGTGGGAAGGCCAGAGTGGCACTGCAAG-----GAGATGATGCTGATGCTCTTCATCT 617  
QY 473 GTATCAGAGTGAGAGATCCCTAGGGGATACCTCCCTTCAGTCACCAACAGCCAGATGTC 532  
DB 618 GCTCCGATGACCCACATCATCGGATTCACGCCCTGACAAAGCGGTCAA-----CAGATATC 673  
QY 533 AGGAGAAATCCCTGGAGCTTCCCTGATCATCATCATGATGTCCTTACGGGGAAACATCAG 592  
DB 674 TGGAGAGATTCCTGATGCTTCCCTGACCGTCACTTATCCGACAGTCCCAACATCGAGCTA 733  
QY 593 CAGAGGTGCTCGTTCCTTATGTAATCAATTCACCAATCCGTCAGGGAGTTC---TC 649  
DB 734 TGTGATCCAAGCGTCCAGTTCCTGTGAGGATTTGGACCCCTCGAAGGACTTGAATTC 793  
QY 650 CGGCAGTATTGGGAATGTTGCTTGGAAAGAGAGAGTGTGATGCTGGAAATCAAGCAGGA 709  
DB 794 CTATGGCTTAATAGTGTGACTGGAAGGAAGAGTGTGAGAGCTGGAGGGTTAAACAGGA 853

QY 710 CAAGGCTGGCATTCCTCCATGACT 769  
DB 854 CAAAATATGATCAAGTGAATGACT 913  
QY 770 TACTGACATCGATGATCACTGACT 829  
DB 914 GACTGCTCAAAATGGA----- 955  
QY 830 TCGCCAGCGCTCATCTAGAAAT 889  
DB 956 AGGCTTACCTTTGAGCCGATGATG 1015  
QY 890 GGTCAATGTTCTCGCGTGTGTT 949  
DB 1016 AGTGATCATTCCTCCGCTTATG 1075  
QY 950 TGTGCGTATGATATACCCACTG 1009  
DB 1076 AGTGCCTGATGCTTATGGAAT 1135  
QY 1010 ATCCTGATCTGATGATCAGTTC 1069  
DB 1136 GTCTTGGCTTCTAGATCAGTTC 1195  
QY 1070 TAGACTGGCTTTAAGGTATGAG 1129  
DB 1196 CAGGCTTGCATGAGGTATGAT 1255  
QY 1130 ATTTGCTAGTACAGTGCAGCC 1189  
DB 1256 CTTCGCTAGTACAGTGGATCCA 1315  
QY 1190 ATCCATCTTGGCTTGTGATTAT 1249  
DB 1316 GTCCATCTTCTGTGGATTAG 1375  
QY 1250 AGCTCAATGCTGACTTTTGG 1309  
DB 1376 TTCAAGTATGCTGACTTTTGG 1435  
QY 1310 ACCATTGCTGAAGAACTATGAC 1369  
DB 1436 TCCCTTTTGAAGAACACAA 1495  
QY 1370 AATTGATTACCTGAAAGACAA 1429  
DB 1496 ATAGATTAACCTGAAGACAA 1555  
QY 1430 GAGAGAAATGAAGAAATTTAA 1489  
DB 1556 GAGGGAGTATGAAGAAATTCNA 1615  
QY 1490 CCCCAGGAGGATGGATCATG 1549  
DB 1616 GCTTGAAGAGGGTGGACCATG 1675  
QY 1550 TCATCTGGAATGATTCAGGTT 1609  
DB 1676 CCATCTGCGATGATTCAGGTT 1735  
QY 1610 TGAGCTCCCGCTTTAGTTTAT 1669  
DB 1736 TGAGTTTACCACGCTTGTCTAT 1795  
QY 1670 GAAGGCTGTGCCATGAATGCC 1729  
DB 1796 GAAGGCTGTGCAATGAATGCC 1855  
QY 1730 CATGTTGAATCTTGTGATGAT 1789  
DB 1856 TCTTCTCAATGATGATTCGAC 1915

Qy	1790	GTGCTTCCTAATGGATCCAAACCTTAGTCCGCAAGTCGTATTATGTCAGTTCGCCACAAG	1844
Db	1916	GTGCTTCATGATGGATCCGGCTCTTAGAAGGAAACACTGTGTATTGTACAATTTCCACAG	1975
Qy	1850	GTTTGATGGANTGATAGAATGATCGATATGCAACAGGAAACACTCTCTTTTTCATAT	1909
Db	1976	ATTTGATGGCATTTGACAGATCGATATGCTAATCGGAACATAGTTCTTCCTTGATAT	2035
Qy	1910	TAACCTTGGGGCCCTTGACGGCATTCAGAGCACCACTGTTTATCTGGGAACCTGTTGTGTTT	1969
Db	2036	CAACATGAAGGCTCGGATGGCATTCAGGTCACCTGTTTACGTGGGAACAGATGCTGCTTT	2095
Qy	1970	CACAGAACAGCTATCTATGTTTATGACC---CCCAATTAAAGCGGAAGAACCAAGTTT	2026
Db	2096	CAATAGCAGGCTTTGATTTGGATACGATGCTGTTTGTAGCTGAAGCTGATCTGGAGCCAA	2155
Qy	2027	CTTTGGCATCACTATGTGGGGGCAAGAGCGCAGCAAGTCAAGAAAGGAGGCTCAGA	2086
Db	2156	CATTGTTTATTAAGAGCTGCTGTGGTAGAAGGAAGAAAGCAAGAGTTATATGGATAG	2215
Qy	2087	TAAGAAAAAGTCGAACAAGCATGTGGACAGTCTCTTCCAGTATTTCAATCTCGAAGACAT	2146
Db	2216	TCAAAGCGTATTATCAAGAGAACAGATCTTCAGCTCCCATCTCAATATGAGNAGCAT	2275
Qy	2147	AGAGGAGGTTGTAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAAT	2206
Db	2276	CGAAGAGGTTATTGAAGT-----TACGAGGATGAAGGTCAGTGTATTATGCCCAGAG	2329
Qy	2207	GACCTTAGAGAGAGATTTGGCCAGTCAAGCAGCATTTGTTGCCCTCCACTCTGATGGNATA	2266
Db	2330	GAAATTTGGGAAGACGCTTTGGTCAGTCTCCTATTTTCAATGCTATCCACCTTTATGACACA	2389
Qy	2267	TGTTGTTGTTCTTCAGTCTCCACTCCAGATCTCTTTTGAAGAAGCTATCCATGTTCAT	2326
Db	2390	AGTGGCATACCACCTTCAACAAACCCAGCTCTCTACTAAGAGNAGCTATCCATGTTCAT	2449
Qy	2327	AAGTTGTGGCTATGAGGACAAGTCTGAATGGGGAACTGAGATTGTTGGATCTATGGATC	2386
Db	2450	CAGTTGTGGATATGAGGACAAAACCTGAATGGGGAAGAGATTTGGCTGGATCTATGGTTC	2509
Qy	2387	TGTCACAGAGATATTCTTTACTGGATTCAAGATGCACCAAGAGGCTGGCGTTCAGTCTA	2446
Db	2510	AGTAAGGAGGATTTCTCACTGGGTTTAAATGCTCAAGGGGCTGGCAATCAATCTA	2569
Qy	2447	TTGCATGCCCAAGCGCCAGCTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATCGTCT	2506
Db	2570	CTGCATGCCACACGACCTTGTTCAGGGTCTCGCACCATCAATCTTTCCGATCGTCT	2629
Qy	2507	GAACCAAGTGCTCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGCGGCAATCGCC	2566
Db	2630	TAATCAGGTGCTCCGTTGGGCTCTTTGGGTCACTGGGAAATTTCTGCTTAGTAGACATTTGC	2689
Qy	2567	CTTATGTTATGGCTACGGAGGGCGCTCAAGTTCTTGAGAGATTCGCTTACATCAACAC	2626
Db	2690	TATCTGGTATGTTTCAATGGACGATGAAGCTTTTGAGAGGCTGGCTTACATCAACAC	2749
Qy	2627	CACCATTTACCACACTAACCTCTCTCCGCTTCTAGCTATTGTTATTTGCTCGCTATCTG	2686
Db	2750	TATTGTATATCCAATCACATCCATTCGCTTATTGCCCTATTGTGTGCTTCCCGTATCTG	2809
Qy	2687	TCTGCTCACTGGAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCAT	2746
Db	2810	CCTCCTTACCAATAAATTTATCATCTCGAGATTAGCAATTTATGCTGGGATGTTCTTCAT	2869
Qy	2747	TGGCTCTTCTCTTCAATTTTGGCCACTGGTATCTTGGATGAGGTGGATGGTGTGG	2806
Db	2870	TCTTCTTTTCCGCTCCATTTTGGCCACTGGTATATTGGAGCTTAGATGGAGTGGTGTGG	2929
Qy	2807	CATTGACGAGTGGTGGAGGAATCAACAGTTCCTGGGTCAATGGAGGTATCTCTGCACATCT	2866
Db	2930	CATTGAAGATTTGGTGAGAAATGAGCAGTCTTTGGGTATTGTTGGTGCACCTCTGCCCATCT	2989
Qy	2867	GTTTGGCGTCTTTTCAGGGTCTTCTGAAGGTGCTTGGCCGTTATCGACACCAACTTCACTGT	2926

[illegible]

Search completed: February 25, 2003, 18:32:22  
Job time : 730 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 15:04:05 ; Search time 4739 Seconds  
(without alignments)  
12391.820 Million cell updates/sec

Title: US-09-900-237-29  
Perfect score: 3626  
Sequence: 1 gcacgagaaccccgctcca.....tcacatttgaggaggtttt 3626

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- EST:\*
- 1: em\_estba.\*
  - 2: em\_esthum.\*
  - 3: em\_estin.\*
  - 4: em\_estmu.\*
  - 5: em\_estov.\*
  - 6: em\_estpl.\*
  - 7: em\_estro.\*
  - 8: em\_htc.\*
  - 9: gb\_est1.\*
  - 10: gb\_est2.\*
  - 11: gb\_htc.\*
  - 12: gb\_est3.\*
  - 13: gb\_est4.\*
  - 14: gb\_est5.\*
  - 15: em\_estfun.\*
  - 16: em\_estom.\*
  - 17: gb\_gss.\*
  - 18: em\_gss\_hum.\*
  - 19: em\_gss\_inv.\*
  - 20: em\_gss\_pln.\*
  - 21: em\_gss\_vrt.\*
  - 22: em\_gss\_fun.\*
  - 23: em\_gss\_mam.\*
  - 24: em\_gss\_mus.\*
  - 25: em\_gss\_other.\*
  - 26: em\_gss\_pro.\*
  - 27: em\_gss\_rod.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2572.2	70.9	3763	11	AV108113 Zea mays
2	2402	66.2	3728	11	AV112236 Zea mays
3	2199	60.6	3696	11	AV110079 Zea mays
4	1373.2	37.9	3898	11	AV110415 Zea mays
5	1320.4	36.4	2872	11	AV104236 Zea mays
6	1319.8	36.4	3788	11	AV103701 Zea mays

7	1315.6	36.3	3783	11	AV104730 Zea mays
8	1267.8	35.0	3897	11	AV103655 Zea mays
9	762.2	21.0	880	14	BM816138 HCL09B12
10	751.6	20.7	916	14	BM816139 HCL05G09
11	750.4	20.7	762	14	BQ802778 WHE2829_H
12	729.4	20.1	870	12	BG368813 HVSME1002
13	689.6	19.0	767	14	BQ578769 WHE0309_E
14	689.4	19.0	710	14	BQ282057 WHE3028_D
15	687.8	19.0	750	12	BF624748 HVSME001
16	684.2	18.9	712	14	BQ620630 TALR1138D
17	683.8	18.9	978	12	BG321258 Zm04_03F1
18	668.6	18.4	755	12	BF259426 HVSMEf001
19	663	18.3	705	9	AJ433167 AJ433167
20	651.4	18.0	659	13	BJ214460 BJ214460
21	649.6	17.9	678	9	AJ433236 AJ433236
22	644.4	17.8	659	14	BQ620294 TALR1173A
23	637.6	17.6	676	14	BQ468579 HMO1K12T
24	630	17.4	654	10	AV914933 AV914933
25	628.4	17.3	661	13	BJ277746 BJ277746
26	628	17.3	656	14	BQ620013 TALR1138D
27	616.6	17.0	627	13	BJ287480 BJ287480
28	615.6	17.0	694	13	BJ282938 BJ282938
29	614.8	17.0	634	13	BJ280838 BJ280838
30	611.2	16.9	625	13	BJ314568 BJ314568
31	609.2	16.8	630	13	BJ317016 BJ317016
32	601.6	16.6	671	9	AL504419 AL504419
33	600.2	16.6	680	10	AV920051 AV920051
34	592.8	16.3	672	13	BJ280905 BJ280905
35	591.6	16.3	643	14	BQ465284 H003C22T
36	588.8	16.2	625	14	BQ471177 HVO1F24T
37	581.4	16.0	603	10	BE403558 WHE0434_A
38	580.6	16.0	615	10	AV932043 AV932043
39	580.4	16.0	669	13	BJ466285 BJ466285
40	578.2	15.9	634	12	BG414971 HVSMEK000
41	577.6	15.9	584	13	BJ233686 BJ233686
42	576	15.9	593	13	BJ320053 BJ320053
43	576	15.9	631	10	AV913277 AV913277
44	571	15.7	583	13	BJ207181 BJ207181
45	571	15.7	681	13	BI959339 HVSMEf001

ALIGNMENTS

RESULT 1  
AV108113  
LOCUS AV108113 3763 bp mRNA linear HTC 25-MAY-2002  
DEFINITION Zea mays PCOL26465 mRNA sequence.  
ACCESSION AV108113  
VERSION AV108113.1 GI:21211191  
KEYWORDS HTC.  
SOURCE Zea mays.  
ORGANISM Zea mays

REFERENCE  
AUTHORS Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Archer,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
TITLE Zea Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 3763)  
AUTHORS Coe,E.C.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA

FEATURES  
Source  
1. 3763  
/organism="Zea mays"  
/db\_xref="MaizeDB:637367"  
/db\_xref="taxon:4577"  
/clone="PCOL26465"

/clone_lib="Maize Mapping Project/DuPont Consensus Library"									
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"									
BASE COUNT	863 a	984 c	1011 g	904 t	1 others				
ORIGIN									
Query Match	70.9%; Score 2572.2; DB 11; Length 3763;								
Best Local Similarity	86.4%; Pred. No. 0;								
Matches 2864; Conservative 1; Mismatches 439; Indels 9; Gaps 2;									
QY	4	CGAGGAACCCCGCTCCAGCTCTGTCGTCGCGGTGGATCGCTCTGCCGCGCATGG	63						
Db	282	CCAGGCCCAAGCTCCAGCGCAGCTCCCTCGACCTTTCTCGCGAGCTCGCTTGCCATGG	341						
QY	64	ACGGCGACGCGACGCCCTGAAGTCCGGGAGGACGGCGGGGAGCTGTGCCAGATCT	123						
Db	342	AGGCGACGCGACCGCGCTGAAGTCCGGGAGGCGCGGTGGCGACAGTGTGCCAGATCT	401						
QY	124	GCGCCAGCGGCTGGGCAACACAGCTTGGACGCGACGCTTTCACCGCCTGCGACGTGCC	183						
Db	402	GCGGCGACGGCTGGGCAACACACGCGGAGGGGAGCTTTCGCGCCTCGACGCTCGG	461						
QY	184	GCTTCCCGTCTGCGCGCCCTGCTACGAGCAGCAGCGCAGGAGGGCACCCAGCCTGCC	243						
Db	462	GGTTTCGGGTGTCGCGCCCTGCTACGAGTACGAGCGCAGGACGCGACGCGGCTGCC	521						
QY	244	TCCAGTCAAGACCAAGTACAAGCGCCACAGAGGGAGCCCGCGATCCGCGGGAGGAAG	303						
Db	522	CCCAGTCAAGACCAAGTACAAGCGCCACAAAGGGACCCCGCGATCCGTGGGGAGGA	581						
QY	304	GCGACGACTGATCCCGATGATGGTAGTACCTTCACTACCTGCTGCTGCGACGTGAGG	363						
Db	582	GAGCGACACTGATGCCGA-----TAGCGACTTCAATTACCTTGGATTCGCAATGAGG	635						
QY	364	ACCAGAAGCAGAAGATTGCTGACAGGATGCGCAGCTGGCGATGAACACCGGGGGCAGTG	423						
Db	636	ACCAGAAGCAGAAGATTGCCGACAGATGCGCAGCTGGCGATGAACGTTGGGGCAGCG	695						
QY	424	GCATGTTGGCCACCCCAAGTATGACAGTGGCGAGATGGCCCTCTCCAAGTATGACAGTG	483						
Db	696	GGGATGTTGGTCGCCCAAGTATGACAGTGGCGAGATGGGGCTTACCAAGTATGACAGTG	755						
QY	484	GAGAGATCCCTAGGGGATACGCTTCCAGTCAACCAACAGCCAGATGTACAGAGAATCC	543						
Db	756	GCGAGATTCTCGGGATACATCCCATGCTACACGCCAGATCTCAGGAGAAATCC	815						
QY	544	CTGGAGCTTCGCTGATCATCATATGATGTCCTTACGGGAAACATCAGCAGACGTGCTC	603						
Db	816	CTGGTGTCTCCCTGACCATCATATGATGTCCTTACGGGAAACATTGGCAAGCGTCTC	875						
QY	604	CGTTTCCCTATGTAATCATTCACCAATCCGTCAAGGGAGTTCTCCGCGAGTATTGGGA	663						
Db	876	CATTTCCCTATGTAACCATTCGCCAAATCCGTCAAGGGAGTTCTCTGGTAGCATTTGGGA	935						
QY	664	ATGTTGCCCTGGAAGAGAGAGTTGATGGCTGGAAATGAAGCAGGACAAAGGTCGCTTC	723						
Db	936	ATGTTGCCCTGGAAGAGAGAGTTGATGGCTGGAAATGAAGCAGGACAAAGGACGATTC	995						
QY	724	CCATGACTAATGGGACAGCAATTCCTCCTGTAAGTCCGGGAGCTACTGACATCATGATG	783						
Db	996	CCATGACGAATGGGACAGCAATTCCTCCTGTAAGTCCGGGAGCTACTGACATCATGATG	1055						
QY	784	CATCTACTGAATACAACTGAAGACGCTTTACTGAATGATGAACCTCGCCAGCCTCTAT	843						
Db	1056	CATCACTGATTTACAACATGGAGATGCCCTTATTTGAAGCAGGAACCTCGACAGCCTCTAT	1115						
QY	844	CTAGAAAAGTCCCCATTCCTTCTCCAAAATAAATCCCTACAGAATGGTCATTTGTTCTGC	903						

Db	1116	CTAGGAAAGTTCCACTTCCCTT	1175						
Qy	904	GGTTGGTTGTTCTAAGCATCTT	963						
Db	1176	GATTGATTGTTCTAAGCATCTT	1235						
Qy	964	ACCCACTGTGCTTTTATCTGTT	1023						
Db	1236	ACCATATATGGCTTCTATCTG	1295						
Qy	1024	ATCAGTTCCGAAAGTGGTTTC	1083						
Db	1296	ATCAGTTCCGAAAGTGGTTTC	1355						
Qy	1084	GGTATGACCCAGAGAGTGAAC	1143						
Db	1356	GGTATGACCCAGAGAGTGAAC	1415						
Qy	1144	TCGACCCCTTTGAAGGAGCCAC	1203						
Db	1416	TCGACCCCAATGAAGAGCCTC	1475						
Qy	1204	TTGATTATCCCGTGGCAAGGT	1263						
Db	1476	TGGATTACCCGTGGATAAGGT	1535						
Qy	1264	CTTTTGACGCAATGGCTGAGAC	1323						
Db	1536	CATTGATGCACTAGCTGAGAC	1595						
Qy	1324	AGTATGACATTTGAACCCAGAG	1383						
Db	1596	AGTACAACTTTGAACCTTAGAG	1655						
Qy	1384	AAGACAAAGTCCAGCCTTCAT	1443						
Db	1656	AGCAAAAGTGCACCTTCAT	1715						
Qy	1444	AATTTAAATCAGGATAAATGC	1503						
Db	1716	AATTTAAAGTTAGGTAATG	1775						
Qy	1504	GGATCATGAAGATGGCACAC	1563						
Db	1776	GGATCATGAAGATGGCACAC	1835						
Qy	1564	TTGAGTTTCTTCTTGGTTCAC	1623						
Db	1836	TTGAGTTTCTTCTTGGTTCAC	1895						
Qy	1624	TAGTTATGTCCTCTGGAAT	1683						
Db	1896	TGCTCATGTTCTCTGGAAT	1955						
Qy	1684	TGAATGCCCTTGTCTGTCCTC	1743						
Db	1956	TGAATGCTCTGTTCTGTCCTC	2015						
Qy	1744	ATTGTGATCACTACATCAAA	1803						
Db	2016	ATTGTGATCACTACATCAAA	2075						
Qy	1804	ATCCAACTAGCTCCGCAAGT	1863						
Db	2076	ACCTTAACCTAGGAAGAGTGT	2135						
Qy	1864	ATAGGAATGATCGATGCAAA	1923						
Db	2136	ACAGGAATGATCGATGCAAA	2195						
Qy	1924	TTGACGGCATTCAGGACCAAGT	1983						
Db	2196	TTGATGGCATCCAGGACCAAGT	2255						



QY	1984	TCATGTGGTTATGAGCCCCCAATTAAAGCGGAAGAACCGAGTTCCTTGGCATCACTATGTG	2043
Db	2256	TATATGGTTATGAGCCCCCAATTAAAGCAGAAGAGGGTGTCTTGTCTCATCACTATGTG	2315
QY	2044	GGGCGAAGAAAGCAAGCAAGCTCAAGAAAGAGGCTCAGATAAGAAAGAGTCCAAACA	2103
Db	2316	GGGGTAGGAAGAGCAAGCAAACTCAAGAA--GGGCTGGACAAGAGAGTCCGACA	2372
QY	2104	AGCATGTGACAGTTCTGTTCACAGTATTCAATCTCGAAGACATAGAGAGGGGTGTGAAG	2163
Db	2373	AGCATGTGACAGTTCTGTGCCAGTATTCAACCTTGAAGATATAGAGAGGGAGTTGAAG	2432
QY	2164	GTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTTAGAGAAAGAT	2223
Db	2433	GGCTCGGATTTGACGACGAGAAATCAGTCTTATGTCTCAAAATGAGCTCGAGAAGAGAT	2492
QY	2224	TTGGCCAGTCACGACGATTTGTGTGCTCCACTCTGATGGAAATAGTGTGGTTCTTCAGT	2283
Db	2493	TTGGCCAGTCCGACGGTTTGTGGCTCCACTCTGATGGAGATAGTGTGGTTCTTCAGT	2552
QY	2284	CTCCACTCCAGAAATCTCTTTGAAAGAAAGCTATCCATGTCTAAAGTTGTGGCTATGAGG	2343
Db	2553	CCGCAACTCCGGAGTCTCTTCTGAAAGAAAGCTATCCATGTTTATAAGCTGTGGCTATGAGG	2612
QY	2344	ACAAGTCTGAATGGGAACTGAGATTTGGATTTGGATCTATGGATCTGTGCACAGAAGATATTC	2403
Db	2613	ACAAGACTGAATGGGGAATGAGATCGGGTGGATCTACGGTCTCTGTGCACAGAAGACATTC	2672
QY	2404	TTACTCGATTCAAGATGCACGCAAGAGGCTGGCGTTGAGTCTATGTGATGCCCAAGGCC	2463
Db	2673	TCACCGGATTCAGATGCACGGCGAGGCTGGCGGTGATCTACTGATGCCCAAGGGCG	2732
QY	2464	CAGCTTTCAGGGATCTGCCGCCCACTCAATCTTTCAGATCGCTGGAACCAAGTCTCGCGT	2523
Db	2733	CAGCTTTCAGGGGTCTGCCGCCCACTCAATCTTTGGGACCGTCTGAACACGAGTCTCGGT	2792
QY	2524	GGGCTCTCGGTTCTGTGAAATCTTTTCAGCGCGCATTTGCCCTTATGGTATGGCTACG	2583
Db	2793	GGGCTCTTGGTCCGGTGAGATCTCTTCAGCGCGCATTTGCCCTGTGTAGCGGTACG	2852
QY	2584	GAGGCGCCTCAAGTTCCTGGAGAGATTCGGTTACATCAACACCAACCATTTACCCACTAA	2643
Db	2853	GAGGCGGCTCAAGTTCCTGGAGAGATTCGGGTACATCAACACCAACCATCTACCCGCTCA	2912
QY	2644	CTCTCTCCGGCTCTAGTCTATTGTATATGCGCTGATCTGTCTGCTCACTGGAAAGT	2703
Db	2913	CGTCCATCCGGTCTCTCATCTACTGATCCTGCCCGCATCTGTCTGTACACGGAAAGT	2972
QY	2704	TCATCATGCCAGAGATTACCAACTTGGCCAGTATCTGGTTTCATTTGGCGCTCTTCTTTCAA	2763
Db	2973	TCATCATTTCCAGAGATCAGCAACTTCGCCAGAGATCTGGTTTCATCTCCTCTTCACTCGA	3032
QY	2764	TTTTGCCCACTGGTATCCTTGAGATCAGGTGGAGTGGTTGGCATTTGACAGTGGTGA	2823
Db	3033	TCCTTCCCGCGGCATCCTGGAGATGAGTGGAGCGGGGTGGGCATCGACAGTGGTGA	3092
QY	2824	GGAAATGAACAGTTCTGGGTCAATTTGGAGGTATCTCTGCAACATCTGTTTGGCGCTCTTTCAGG	2883
Db	3093	GGAAACGAGAGTTCTGGGTGATCGGGGGGCATCTCCGCGCAGCTCTTGGCGGTGTTCAGG	3152
QY	2884	GTCTTCTGAAGGTCTTTCGGGTATCGACACCAACTTCACCTGCACCTCAAAAGGCTAATG	2943
Db	3153	GGCTGTCTAAGGTGTCTGGCGGCATCTGCACCAACTTCACCGTCACTTCGAAGGCTTCGG	3212
QY	2944	ACGAAAGAGCGCACTTTGCTGAGCTCTACATGTTTCAAGTGGACGACAGTTCCTCATCCCTC	3003
Db	3213	ACGAGGACGCGACTTTCGCGGAGCTGCTACATGTTTCAAGTGGACGACGCTCTCTGATCCCG	3272
QY	3004	CGACGACCAATTTGATCATTAACATGGTTGGTGGTCTGTGGCACTCTCTACGCCATCA	3063
Db	3273	CCACCACTCTGTGATCATCAACTGGTTCGGGTCTGGCGGATCTCTCTACGCCATCA	3332

QY	3064	ACAGTGGTTACCAATCATATGGGGCCGCTCTTTGGGAAGCTCTTCTTTGGCCCTTCCTGGGTGA	3123
Db	3333	ACAGCGATACCACTCGCTGGGGCCGCTCTTCGGCAAGCTCTTCTTCGCCCTTCCTGGGTGA	3392
QY	3124	TTGTTCACTTATACCCATTTCTCTCAAGGTCCTTATATGGCGAGGCAAAACCGCACACCGACGA	3183
Db	3393	TCGTCACCTGTACCCGTTCTCAAGGCGCTCATATGGCGAGGCGAGCAACCGCACCCGACCA	3452
QY	3184	TTGTCATCGTGTGGCGTGTCTCTCTCGCTCTTATCTTCTTCTCTGCTGTGGGTTCTGCTGTG	3243
Db	3453	TCGTCGTCTGTGGCCATCTCGCTGGCGTCCATCTTCTCTCTGCTGTGGGTTCTGCGATCG	3512
QY	3244	ATCCATTTCACTACCCGTCCTCGCTGGCCCAATATCCAAACCTGTGGCATCAACTGCTAGG	3303
Db	3513	ACCCCTTACACACCCGCGTCACTGCGCCGGATACCCAGACGCTGTGGCATCAACTGCTAGG	3572
QY	3304	AAAGTGGGAGTTT	3316
Db	3573	GAAGTGGGAAGTT	3585
RESULT	2		
LOCUS	AY112236	3728 bp	linear
DEFINITION	Zea mays CL1160_1 mRNA sequence.		
ACCESSION	AY112236		
VERSION	AY112236.1	GI:21216826	
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitsitt, M.S., Athur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 3728)		
AUTHORS	Coe, E.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
FEATURES	Location/Qualifiers		
source	1..3728		
	/organism="Zea mays"		
	/db_xref="MaizeDB:630049"		
	/db_xref="taxon:4577"		
	/clone="CL1160_1"		
	/clone_lib="Maize Mapping Project/DuPont Consensus Library"		
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
BASE COUNT	782 a	845 c	925 g
ORIGIN	878 t	298 others	
Query Match	66.2%	Score 2402;	DB 11; Length 3728;
Best Local Similarity	80.7%;	Pred. No. 0;	
Matches 2660;	Conservative	0; Mismatches 631;	Indels 4; Gaps 2;
QY	36	CGGTTGGATCGCTCGCGGCCATGACCGCGACGGAGCGCCCTGAAGTCCGGGAGG	95
Db	214	CGGTTGGGCGAGCTCGCTGCCATGAGGCGGACGCGGCGCTGAAGNNNNNNNN	273
QY	96	CACGGGCGGGGACGTGTCCAGATCTGCGCCGACGGCTGGGCACCACTGTGGACGC	155
Db	274	NNNNNNNNNNNNNGTGGCAGATCTGCGCGTGGCTGGGCACTACGCGGAGGGA	333

QY 156 GACGCTTTCACGGCCTGCGAGCTGTCGCCGCTTCCGGCTGTCGCCGCCCTGCTACGAGCAC 215  
DB 334 GACGCTTTCACGGCCTGCGAGCTGTCGCCGCTTCCGGCTGTCGCCGCCCTGCTACGAGTAC 393  
QY 216 GAGCCAAAGGAGGACCCAGCGCTGCTCCAGTCAAGACCAAGTCAACAGCGCCACAGA 275  
DB 394 GAGCCAAAGGAGGACCCAGCGCTGNNNNAGTGCNNNNCAAGTCAACAGCGCCACAG 453  
QY 276 GGGAGCCAGGATCCCGGGGAGGAGGCGACGACACTGATGCCGATGATGGTAGTAC 335  
DB 454 GGGAGTCCAGCGATCCGAGGGGAGGAGGACGACGATGATGCCGATGATGCTAGCGAC 513  
QY 336 TTCAACTACCGTGCATCTGGCACTCAGGACCAAGCAGAGAGAGATTGCTGACAGATGCGC 395  
DB 514 TTCAACTACCGTGCATCTGGCAATGACGACGAGCAGAGAGAGATTGCTGACAGATGCGC 573  
QY 396 AGCTGGCGCATGAACACCCGGGGGAGTGGCAATGTTGGCCACCACCAAGTATGACAGATGGC 455  
DB 574 AGCTGGCGCATGAATGCTNNNNCAGCGGGATGTTGGCCGCCCAAGTATGACAGTGGT 633  
QY 456 GAGATCGGCCCTCCAAAGTATGACAGTGGAGAGATCCCTAGGGATAGCTCCCTTCAGTC 515  
DB 634 GAGATCGGCCCTCCAAAGTATGACAGTGGAGAGATCCCTAGGGATAGCTCCCTTCAGTC 693  
QY 516 ACCAACGCCAGATGTCAGGAGAAATCCCTGGAGCTTCGCCCTGATCATCATGATGTC 575  
DB 694 NNNNNNNNNATTTCCGGGAGAAATCCCTGGTCTTCCCTGACCATATGATGTC 753  
QY 576 CCTAGGGGAACATCAGCAGAGCTGCTCCGTTTCCCTATGTAATCATTCACCAAAATCCG 635  
DB 754 CCTACTGGGAACATTTGGCAGGCGGCCCATTTCCCTATATGAATCATTCATCAAAATCCG 813  
QY 636 TCAAGGAGTCTCCGGCAGTATTGGGAATGTTGCCCTGGAAGAGAGATTGATGGCTGG 695  
DB 814 TCGAGGGAATTTCTGCTAGCGTTGGGAATGTTGCCCTGGAAGAGAGATTGATGGCTGG 873  
QY 696 AAATGAAGCAGGACAGGGTGCATTTCCCATGACTAATGGGCAAGCATTTGCTCCCTCT 755  
DB 874 AAATGAAGCAGGACAGGGTGCATTTCCCATGACTAATGGGCAAGCATTTGCTCCCTCT 933  
QY 756 GAAAGTCCGGCAGCTACTGACATCGATCTACTGAATCAACATCAACATGGAAGAGCTTTA 815  
DB 934 GAGGCGCGGGTGTGGTGATTTGATGATCACTGATTAACATCAACATGGAAGATGCCTTA 993  
QY 816 CTGAATGATGAACATCGCCAGCGCTCTATCTAGAAAAGTCCCATTTGCTTCCCAAATA 875  
DB 994 TTAACGATGAACATCGCCAGCGCTCTATCTAGAAAAGTCCCATTTGCTTCCCAAATA 1053  
QY 876 AATCCCTACAAATCGTGCATTTGCTGGGTTGTTGTTAAGCATCTTCCCTGCACATC 935  
DB 1054 AATCCATACAGATGGTCAATTTGCTAGGATTTGTTTCAAGCATCTTCTTGCACATC 1113  
QY 936 CGTCTCAAAATCGTGCATTTGCTAGTACCATCCACTGTCGCTTTTATCTGTTATGTCAG 995  
DB 1114 CGGATCAAAATCGTGCATTTGCTAGTACCATCCACTGTCGCTTTTATCTGTTATGTCAG 1173  
QY 996 ATTTGGTTGCTTTTCTGATACTGATCAGTTCGGAAGTGGTTCCCATCAACCGG 1055  
DB 1174 ATCTGGTTGCTTTTCTGATATTTGATCAGTTTCCAAAGTGGTTCCAAATCAACCGG 1233  
QY 1056 GAGACCTACCTTGTAGACTGGCTTTAAGGTATGACCGAAGGTGAACCTCTCAGTTG 1115  
DB 1234 GAGACTTACCTTGTAGACTGCGATTAAGGTATGACCGAAGGTGAGCCATCTCAGTTG 1293  
QY 1116 CTGCTGTGTACATATTTGTCAGTACAGTGCACCCCTTGAAGGAGCCACCTATCGTCACT 1175  
DB 1294 GCTGCTGTGACANNNGTTCAGTACTGTCGACCCCAATGAAGGAGCCCTCTTGTCACT 1353  
QY 1176 GCCAACACTGTGCTATCCATCTTGTGCTGTTGATTATCCCGTGGCAAGGCTCTTGTCTAT 1235  
DB 1354 GCCAATACCGTGTATCCATCTCGCTGTGACTATCTGTTGGATGAAGGCTCTTGTCTAT 1413  
QY 1236 GTATCTGATGAGGAGCTTCAATGCTGACTTTTGACGCATTTGGCTGAGACTTTCAGATTT 1295

DB 1414 GTATCTGATGATGAGCTGCT 1473  
QY 1296 GCTAGGAAATGGGTACCATTTG 1355  
DB 1474 GCTAGAAATGGGTGCCATTTG 1533  
QY 1356 TACTTTTCCAGAAAATTTGATTT 1415  
DB 1534 TACTTCTCCAGAAAATTTGATTT 1593  
QY 1416 CGCCGGGCCATGAAGAGAGATTT 1475  
DB 1594 CGCCGGGNNNNNNNNNNNNNN 1653  
QY 1476 AAGCATTTGAAGTCCCGGAGG 1535  
DB 1654 AAGCACCNNNNNGTCCCTGAGG 1713  
QY 1536 AACATACCAGGGATCATCTG 1595  
DB 1714 AACAAATACCAGGGACCATCTG 1773  
QY 1596 GATACTCAGGCTAATGAGCTCC 1655  
DB 1774 GATACTCAGGCTAATGAGCTAC 1833  
QY 1656 TTCAGCACCACAAAGAAGGCTG 1715  
DB 1834 TTCAGCATCACAAAGAAGCTG 1893  
QY 1716 ACTAATGGCAATACATGTTGA 1775  
DB 1894 ACCAATGGCAATACATGTTGA 1953  
QY 1776 GTCGAGAAGCTATGCTCTCC 1835  
DB 1954 CTCAGGAAGCTATGCTCTCC 2013  
QY 1836 CAGTTCCCAAAAGGTTTGATG 1895  
DB 2014 CAGTTTCCCAAGAGTTTCGATG 2073  
QY 1896 GTCCTTTTGTATTAATTAATTTGA 1955  
DB 2074 GTGNNNNCGATATTAATTTGA 2133  
QY 1956 ACTGGTTGCTGTTTCAACAGAA 2015  
DB 2134 ACTGGCTGCTGTTTCAACAGAA 2193  
QY 2016 AAGCCAGTTTCTTTGGCATCAC 2075  
DB 2194 AAGGGTGTCTTGTCTATCAC 2252  
QY 2076 AGGAGCTCAGATGAAGAAAAGT 2135  
DB 2253 -GGGCTCAGACAAAGNNNNNG 2310  
QY 2136 CTCGAAGACATAGAGGAGGGTGT 2195  
DB 2311 NNNNNNNNTAGAGGAGGGTGT 2370  
QY 2196 ATGCTCAAAATGAGCTTAGAGAT 2255  
DB 2371 ATGCTCAAAATGAGCTTAGAGAT 2430  
QY 2256 CTGATGAATATGGTGGTGTCC 2315  
DB 2431 CTGATGAATATNNNNNNNNNNNN 2490  
QY 2316 ATCCATGTCAAAAGTTGTGGCTTA 2375

Db 2491 ATCCATGTCATAAGTTGTGGCTACGAGCAAGATGAATGGGAACTGAGATTGGGTGG 2550  
QY 2376 ATCTATGGATCTCTACAGAGATATCTTACTGGATTCAAGATGCACGACGAGCTGG 2435  
Db 2551 ATCTATGGTTCTGTGACGAAGATATCTACATGGTTCAAGATGCACGACGAGCTGG 2610  
QY 2436 CGTTCAGTCTATTGTCATGCCAAGCCGACGCTTTCAAGGGATCTCCGCCATCAATCTT 2495  
Db 2611 CGGTGATCTACTGATGCTTACGCGCCGCTTCAAGGATCGCTCCCATCAATCTC 2670  
QY 2496 TCAGATGCTGTAACCAAGTGTGCGGTGGCTCTCGGTTCTGTGTAATCTTTTCAGC 2555  
Db 2671 TCAGACGCTGTGAACAGAGTGTCTCGGTGGCTCTCGGTTCTAGTGGAAATCTTTTCAGC 2730  
QY 2556 CGGCATTGCCCTTATGATGCTACGAGGCGGCTCAAGTTCCTGGAGAGATTCTGCT 2615  
Db 2731 CGGCATTGNNNNATATGGTACGGGTACGGAGGACGCTGAAGTCTTGGAGAGATTCGCC 2790  
QY 2616 TACATCAACACCACTATTTACCCACTAACTCTCTCCGCTTCTAGTCTATTGTATATTG 2675  
Db 2791 TACATCAACACCACTATTTACCCGCTCAGTCTCCCTCCGCTCTCAATTTACTGTATCTCG 2850  
QY 2676 CTTGCTATCTGTCTGCTACTGGAAGTTCATCATCCAGAGATTAACAACTTGGCCAGT 2735  
Db 2851 CCTGCCATCTGCTGTACGGGGAAGTTTCATCATCCAGAGATCAGCAACTTCGCTAGT 2910  
QY 2736 ATCTGGTTCAATCGGCTTCTTCAATTTTCGCACATGTTATCTCTTGGATGAGGTGG 2795  
Db 2911 ATCTGGTTCACTCTCTCTTCACTTCATCTGATCTTCGCACGCGGTATCTCTGGAGATGAGGTGG 2970  
QY 2796 AGTGGTGTGGCATTGACGAGTGGTGGAGGATGAACAGTTCCTGGGTCAATGGAGGTATC 2855  
Db 2971 AGCGGGTGGGCAATCAGAGTGGTGGAGGACGACAGTTCGGGTCAATCGGAGCATC 3030  
QY 2856 TCTGCACATCTGTTTCCGCTTCTTCAAGGTCTTCTTCAAGGTGCTTCCGCTATTCGACACC 2915  
Db 3031 TCCGCCACCTCTTCCGCTTCTTCCAGGGCTCTCTCAAGGTGCTTCCGCGCATTCGACACC 3090  
QY 2916 AACTTCACCTGTCACCTCAAGGCTATGAGGAAGAGGCGACTTTCCTGAGCTCTACATG 2975  
Db 3091 AACTTCACCTGTCACCTCAAGGCTCGGATGAAGAGGCGGACTTCGCGAGCTGTATCATG 3150  
QY 2976 TTCAAGTGGACGAGCTTCTCATCTCCGACGACCATTTTGATCATTAACATGTTGTTGT 3035  
Db 3151 TTCAAGTGGACGACTTCTGATCCCGCCACCATCTCTGATCATCAACCTTGTGCTCGC 3210  
QY 3036 GTGCTGTGCTGGACCTCTACGCCATCAACAGTGTGTTACAAATCATGAGGGCGGCTCTTT 3095  
Db 3211 GTTGTGTCGGCATCTCTACGCCATCAACAGCGGGTACAGTCTGTGGGTCCGCTCTTC 3270  
QY 3096 GGAAGCTCTTCTGCTTCTGGGTGATGTTCACTTATACCATTCCTCAAGGCTCTT 3155  
Db 3271 GGAAGCTCTTCTGCTTCTGGGTGATGTTCACTTATACCATTCCTCAAGGCTCTC 3330  
QY 3156 ATGGGAGGCAAAACCGCACGAGATGTTTCATGCTTGGGCTGCTCTCTCTCGGTTCT 3215  
Db 3331 ATGGGTCGGCAGAACCCGACCCGACCATCTGTTGTTCTGGGAGTCTCTGTTGGGCTG 3390  
QY 3216 ATCTTCTCTTGTGTGGTGTCTGTGTGATTCATTCACCTACCGGCTCTCGCTGGCCAAAT 3275  
Db 3391 ATCTTCTCTTGTGTGGTGTCTGATCGATCCGTTTCAACCAACCGGCTCTACGCGCGGAT 3450  
QY 3276 ATCCAACTCTGGCATCACTGCTAGGAAGTGGGA-GTTTGTAGACAGAGAA 3329  
Db 3451 ACTCGAAGCTGTGGCATCACTGCTAGGGAGGTGGGAAGTTGTTAGAAAACAGAGA 3505

RESULT 3  
AY110079  
LOCUS AY110079 3696 bp mRNA linear HTC 25-MAY-2002  
DEFINITION Zea mays Cull164\_1 mRNA sequence.  
ACCESSION AY110079  
VERSION AY110079.1 GI:21214162

KEYWORDS HTC.  
SOURCE Zea mays.  
ORGANISM Zea mays.  
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.  
AUTHORS 1 (bases 1 to 3696)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and ringey,S.V. Design of Zea Mapping Project/DuPont Consensus Sequences for the Design of Overgo Probes  
JOURNAL Unpublished (2002)  
REFERENCE 2 (bases 1 to 3696)  
AUTHORS Coe,E.C.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA  
FEATURES  
source  
1. 3696  
/organism="Zea mays"  
/db\_xref="MaizeDB:630055"  
/db\_xref="taxon:4577"  
/clone="Cull164\_1"  
/clone\_lib="Maize Mapping Project/DuPont Consensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project."  
BASE COUNT 831 a 766 c 880 g 945 t 274 others  
ORIGIN  
Query Match 60.6%; Score 2199; DB 11; Length 3696;  
Best Local Similarity 78.6%; Pred. No. 0;  
Matches 2569; Conservative 0; Mismatches 686; Indels 12; Gaps 4;  
QY 38 GGTGGATCGCTCTCCCGCGCATGGACGGGACGGGACGGCCCTGAAGTTCGGGAGCGA 97  
Db 232 GGAGGGCTCGGATTTGGTGGCCATGGACGGCGGC---GACGCCACGAATTCGGGGAAGCA 288  
QY 98 CGGGCCGGGACGAGTGTCCAGATCTGCGCGACGGCTGGGACGACCGCTTGGAGGGGGA 157  
Db 289 TGTGCCCGGGACAGTGTGCCAGATCTGCGGACGCGGCTGGGACCGCGGGAGCGGCA 348  
QY 158 CTTCTTACCGGCTCGGACGCTGTCGCGCTTCCCGCTCTGCGCCCGCTCTACGAGCAGA 217  
Db 349 CTTCTTACCGGCTCGGACGCTGTCGCGCTTCCCGCTGTCGCGCCCATCTACGAGTAGCA 408  
QY 218 GCGCAAGGAGGCGCACCGAGGCTGCTCCAGTGCAGACCAAGTACAGGCCACAGAGG 277  
Db 409 GCGCAAGGAGGCGCACCGAGGCTGCGCGAGTGCAGGACTAAGTACAAAGCCCAAGAG 468  
QY 278 GAGCCACGAGTACCAGGGGAGGAGGACGACACTGATCCGATGATGTTAGTGACTT 337  
Db 469 GAGCCACGAGTACACGAGTGGAGAAATGAGGATGTGGATGCTGACGATGTGATGACTA 528  
QY 338 CAATTCACCTGCTATGCGACTGAGGACCAAGACAGAGATGCTGACAGGATGCGGAG 397  
Db 529 CAATTCACCAAGCATCTGGCAACAGGATCAGAAGCAAAAGATTGCTGAGAGAAATGCTCAC 588  
QY 398 CTGGGCGCATGAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGA 457  
Db 589 TTGGCGGACAAACTCACGTGGCAGTG---ATATTGGCCCTGGCTAAGTATGACAGCGGTGA 645  
QY 458 GATCGGCTCTCCAAGTATGACAGTGGAGATCCCTAGGGGATACGCTCCCTTCAGTCTAC 517  
Db 646 AATTGGGCGATGGGAAGTATGACAGTGTGAGATCCCTCGTGGATATATCCCGCTCACTAAC 705  
QY 518 CAACAGCCAGATGTCCAGGAGAAATCCCTGGAGCTTTCGCTGATCATCATCATGATGTCGCC 577  
Db 706 TCATAGCCAGATCTCAGGAGAGATTCTCTGGAGCTTCCCTGATCATATGATGATGCTCTCTGT 765

QY	578	TACGGGAAACATCAGCAGAGTGC	TCGCTCCGTTTCCCTATGTG	GAATCATTCACCAAAATCCGTC	637
Db	766	TGGGAAACATTTGGCAGCGGTGG	ACATCAATTTCTTATGTAAAT	CAATCTCCAAACCCCATC	825
QY	638	AAGGGAGCTTCTCCGGCAGTAT	TGGGAATGTCCTCGTGGAA	GAGAGAGTGTGATGCTGGAA	697
Db	826	GAGGGAGTCTCCGGTAGCTTGG	CAATGTCATGTGGAAGAGAG	GAGGGTGGATGGATGGAA	885
QY	698	AATGAACAGGACAAGGTCGAT	TCCCATGACTAATGGGACA	AGCATTTGCTCCCTCTGA	757
Db	886	AATGA---AGGATAAAGGTGCA	ATTCCTATGACCAATGAA	CAAGCATTTGCTCCATCAGA	942
QY	758	AGTCTGGCGAGCTACTGCAT	CGATCGACTCTCAATACAC	ATGGAAGAGCGCTTTACT	817
Db	943	AGGCGTGGAGTTCGTGAT	TATGATGCTTCTACTGAT	TATAACATGGAAGATGCTTACT	1002
QY	818	GAATGATGAACCTCGCCAGC	CTTACTAGAAAAAGTCC	CCCATTTGCTTCTCCAAAAATAA	877
Db	1003	GAATGATGAACCTCGCAAC	CTTACTAGAAAAAGTG	CCCATTTCTTCCATCCGAATATAA	1062
QY	878	TCCCTACAGAAATGGTCAT	TGTCTCGGGTTGGTTGTT	CTTAAGCACTTCCTGCACTACCG	937
Db	1063	TCCGTACAGAAATGGTCAT	TGTCTAGCTTTGGCTGTT	CTATGCAATATTTCTTGGCCTACCG	1122
QY	938	TCTCACAAATCTGTGCGTA	ATGCATACCACCTGTGG	CTTTTACTGTATATCTGAGAT	997
Db	1123	TATCACACATCTGTGA	CAATGCATCCACTGTGG	CTTTTATCTCCGTCATATGTGAGAT	1182
QY	998	TTGGTTTGTCTTATCTCG	GATACTTGGATTCGCCG	AGTGGTTTCCAAATCAACCGGGA	1057
Db	1183	CTGGTTTGTCTTCTCG	TGATTTTGSATCAGT	TTCCCAAGTGGTCCCAATCAACCGTGA	1242
QY	1058	GACCTACCTTGATAGACT	GGCTTTAAGGTATGAC	CGAGAAGGTGAACGGTCTCAGTTGGC	1117
Db	1243	AACATACCTTGATAGACT	GGCTTTAAGGTATGAC	CGAGAAGGTGAACCATCTCAATTAGC	1302
QY	1118	TGCTGTTGACATATTTG	CTCAGTACAGTCACCC	CTTTGAAGGAGCCACTATCGTCACTGC	1177
Db	1303	TCCGTGTTGATANNNGT	CTCAGTACTGTGAT	CCAAATGAAGGAGCCTCCTCTGTGCTACTGC	1362
QY	1178	CAACACTGTGCTATTC	ATCTTCTGCTTGAT	TATCCCGTGGACAAGGTCTCTTCTGCTATGT	1237
Db	1363	AAATACTGTCTTCCAT	CTTGTCTCGAT	TATCCGGTTGACAAGGTATCTTCTGCTATGT	1422
QY	1238	ATCTGATGACGGAGCT	TTCAATGCTGACTTTT	GACCGCATTTGGCTGAGACTTCAGAGTTTGC	1297
Db	1423	TTCCGATGATGAGCT	GTCTATGCTACTTTT	GATGCTCTCTCTGAAACTTCAGAGTTTGC	1482
QY	1298	TAGAATGGGTACCATTT	GTGAAGAAGTATGAC	ATTGAACCCAGAGCTCCCGAGTTTGA	1357
Db	1483	TAGAAATGGTTCCTG	TATAGAAGTATCAAC	ATAGAGCCTTANGGCCCGCGAATGGTA	1542
QY	1358	CTTTTTCAGAAAAATT	GATTACCTGAAAGACA	AAAGTCCAGCGCTTCATTTGTTAAAGACCG	1417
Db	1543	CTTTTGTCTAGAAAA	ATTGATTCTGAAAGACA	AAAGTCTCAAACCTCATTTGTGGAAGAACG	1602
QY	1418	CCGGGCCATGAAGAGAA	TATGAAGATTTAAAT	CACAGATAAATGCCCTAGTTCCTAA	1477
Db	1603	CCGGGNNNNNNNNNN	NNNNNNNNNNNNNN	NNNCAAGTTCGTATCAATGGTCTGTGAGCAA	1662
QY	1478	GGCATTGAAAGTCCCG	GAGGAAGGTGATCAT	GCAATGSCACACCATGCGCCAGGAAA	1537
Db	1663	GGCACNNNNNGT	TCCCGAGAGGAGT	GATCATGCAAGATGGTACACTTGGCCCTGGGAA	1722
QY	1538	CAATACACAGGATCAT	CTCTGGAAATGATTC	CAGGTTTTCCTTGGTCCACAGTGGTGGCCTTGA	1597
Db	1723	CAATACTAGGACCAT	CTCTGGAAATGATTC	CAGGTTTTCCTGGTCCACAGTGGAGGGCTTGA	1782
QY	1598	TACTGAGGTTAATG	AGCTCCCCCGTTT	AGTTTATGTGTCGTGAAAGCGTCTCGGGTT	1657
Db	1783	CGTTGAAGCAATG	AACTTCTCTGGTT	TATGTGTCGTCTGCTGNNNNNGTCTCGGAT	1842

QY	1658	CGAGCACCACAAAGAGCGCTGGT	TTGAATGCCCTTGTTCGTGTCCTCAGCTGTCCTTAC	1717
Db	1843	CCAACTATCAACAAGAGCGCTGGT	TTGAATGCCCTTGTTCGTGTCCTCAGCTGTCCTTAC	1902
QY	1718	TAATGACAAATACATGTTGAAT	ATTGTGATCACTACATCAACAACAGCAAGCGCTGT	1777
Db	1903	TAATGGCAATACATGTTGAAT	ATTGTGACCACTACATCAATTAATAGCAAGGCTCT	1962
QY	1778	CCGAGAAGCTATGTGCTTCCTTA	ATCCAAACCTAGGTCCGCAAGTCTGTTATGTGCA	1837
Db	1963	TCGAGAAGCTATGTGCTTCCTTA	ATCCAAACCTAGGAAGGAATGCTGTTATGTGCCA	2022
QY	1838	GTTCACCAAAAGGTTTGTATGGG	ATAGGAATGATCGATATGCAAAACAGGAACACTGT	1897
Db	2023	ATTTCCTCAGAGGTTTGTATGGT	ATAGGAATGACCGATATGCAAAACAGGAACACTGT	2082
QY	1898	CTTTTTTGTATTAACCTTGAGG	TTGACGGCATTCAAGGACCAGCTTTATGTGGGAAC	1957
Db	2083	GNNNNCGATATTAACTTTGAGG	TTGACGGCATTCAGGGCCAGCTTTATGTGGGAAC	2142
QY	1958	TGGTGTGTTTTCAACAGAACTTA	TCTATGTTTATGAGCCCCCAATTAAGCGCAAGAA	2017
Db	2143	TGGTGTGTTTTCAACAGAACTTA	TATATGTTTATGAGCCTCCAGCTCAAGNNNNNNN	2202
QY	2018	GCCAGGTTTCTTGGCATCACTTA	GGGCAAGAAGAAGGCAAGCAAGTCAAAAGAAAG	2077
Db	2203	GCCAGGTTTCTTCTCTTCGGT	NNNNAAGNNNNNACGCTCAAAATCTAAGAA---	2259
QY	2078	GAGCTCAGATAAGAAAAGTCG	AGCATGTGGACAGTCTGTTCAGCAATTCATCT	2137
Db	2260	GAGCTCGGAAAAGAAGATCA	GACAGGACAGCTCTGTACCACTATTAACTCT	2319
QY	2138	CGAAGACATAGAGGAGGCTGTT	GTGCTGGGTTTGATGATGAGAAATCAGTCTCAT	2197
Db	2320	CGAAGATATAGAGGAGGATTT	GTCTCAGTTTGATGATGAGAAATCGCTGATTAT	2379
QY	2198	GTCTCAAAATGAGCTTAGAAG	TTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCT	2257
Db	2380	GTCTCAAAATGAGCTTAGAAG	TTGGCCAGTCAGTGNNNNNGTAGCCTCTACTCT	2439
QY	2258	GATGGAATATGGGTGTCTCT	CTCTCCACTCCAGATCTCTTTTGAAGAAGCTAT	2317
Db	2440	GATGGAATATGGGTGTCTCTCA	CTCGCACTCCNNNNNNNNNNNNNNNNNNNNNN	2499
QY	2318	CGATGTCATAAGTTGTGGCTAT	CAAGTCTGAATGGGAACCTGAGATTTGGTTGGAT	2377
Db	2500	TGATGTCATCAGCTGTGGCTAT	CAAAAACCTGACTGGGAACCTGAGATTTGGTGGAT	2559
QY	2378	CTATGGATCTGTACAGAAGAT	TTACTGGATTCAAGATGCACGCAAGAGGCTGGCG	2437
Db	2560	CTATGTTCTGTGTACAGAAG	CTACCGGNNNNNNNNNNNNNNNNNNNNNNNN	2619
QY	2438	TTGAGTCTATTGCAATGCCCAAG	CAGTTTCAAGGATCTGCCCCCACTCAATCTTTC	2497
Db	2620	NNNNNNNNNNNNNNNNNNCTAAG	CAGCTTTCAAGGGATCTGCTCTATCAACCTTTC	2679
QY	2498	AGATCGCTCTGAACCAAGTGTG	GGGCTCTCGGTTCTGTTGAAATCTTTTTTCAGCCG	2557
Db	2680	GGATCGTTTGAATCAAGTGT	GGGCTCTGGTTCATTGAAATCTTTTTTCAGCAG	2739
QY	2558	GCATTGCCCTTATGTATGGG	GAGGGCGCTCAAGTCTCCTCGAGAGATTCGGTTA	2617
Db	2740	GCATTGTCCCATATGATGGG	GAGGGCGGCTTAAATTCCTCGAGAGATTTGCTTA	2799
QY	2618	CATCAACACACCATTTACCCA	CTCTCTCCCGCTTCTAGTCTATTGTAATATGGC	2677
Db	2800	TATCAACACAACAATTTATCCA	CATCAATCCGGCTCTCTCTGACTGCATATTGCC	2859
QY	2678	TGCTATCTCTGCTCACTGGG	TCATCATGCCAGAGATTAGCAACTTGCCAGTAT	2737
Db	2860	AGCAGTTTCTCTCTCACTGGG	TCATCATPCCAAAAGATTAGTAACCTCAGAGAGTGT	2919
QY	2738	CTGGTTTCATTGGGCTCTTCTCT	TTTTTCGCCACTGGTATCTCTTGAGATGAGGTGGAG	2797



Db 1064 AGTCGCTGATGCTATGGAATGTCGGCTAGTATCTGTATCTGCGAGGTCTGGTTGCTT 1123  
QY 1010 ATCTCGATACATGATCAGTTCCGAGTGGTTTCCAAATCAACCGGGAGAGCTACCTTGA 1069  
Db 1124 GTCTGCTGCTAGTACAGTTCCAAATGGTATCCAAATCAACCGTGAGACATATCTCGA 1183  
QY 1070 TAGACTGGCTTTAAGGTATGACCGAGGAAGGTGCAACCGTCTCAGTTGGCTGCTGTGACAT 1129  
Db 1184 CAGGCTTGCACTGAGGTATGATAGAGAGGAGAGCCATCAGAGCTGGCTCCCATGATG 1243  
QY 1130 ATTGTGACATGATGACCCCTTGAGAGAGCCACCTATCGTCACTGCCAACACTGTGCT 1189  
Db 1244 CTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1303  
QY 1190 ATCCATCTGCTGTTGATATCCGTCGACAGAGGTCTCTGCTATGCTATGCTATGATGACGG 1249  
Db 1304 GTCCATCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1363  
QY 1250 AGCTTCAATGCTGACTTTTACGCTGCTGAGACTTTCAGAGTTTGTCTAGGAAATGGT 1309  
Db 1364 CTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423  
QY 1310 ACCATTTGTAAGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1369  
Db 1424 TCCCTTTTGTGAAGACACAAATATGAACCAAGAGCTCCAGAAATTTTACTTTTGTCTC 1483  
QY 1370 AATTGATTACCTGAAAGACAAAGTCCAGCCTTCATTTGTTAAAGACCGCGGCCATGAA 1429  
Db 1484 NNTAGATTACCTGAAGACAAATTCACCTTCATTTGTTAAAGAAAGACCGCAATGAA 1543  
QY 1430 GAGAGAAATGAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTTGAAGT 1489  
Db 1544 GAGGAGTATGAGAAATTCAGAAATGAAGTCAATGCCCTGTTGTCACCAAGCACAGAAAT 1603  
QY 1490 CCCGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549  
Db 1604 GCCTGAAGAGGGTGGACCATGCTGATGGAATGCTTGGCCTGGGAATAAACCTAGGGA 1663  
QY 1550 TCATCTCGGAATGATTCAGGTTTTCCTTGTGTCACAGTGGTGGCTTCTACTAGGGA 1609  
Db 1664 CCATCTGGCATGATTCAGGTGTTCTTGGGGCACAGTGGTGGGCTCAGACATGATGAAA 1723  
QY 1610 TGAGCTCCCGCTTATGTTATGTTGCTGCTGAAAGCGTCCCTGGGTTCCAGCACCAAA 1669  
Db 1724 TGAGTTACCAAGCTTCTGCTATGCTCTGCTGAAAGAGACAGGCTTTCAGCATCAAA 1783  
QY 1670 GAAGGCTGGTGCATGAATGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729  
Db 1784 GAAGGCTGGTGCATGAATGCCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1843  
QY 1730 CATGTTGAATCTTGATGCTGATCAGTACATCAACACAGCAAGGCTGTCGAGAGGCTAT 1789  
Db 1844 TCTTCTCAATGTGATGCTGACCATTAATCAATAGCAGCAAGGCTCTTAGAGAAGCAAT 1903  
QY 1790 GTGCTTCTAATGATCCAAACCTAGTCCGCAAGTCTGTTATGTCAGTTTCCACAAAG 1849  
Db 1904 GTGCTTCTAATGATCCGCTCTAGGAAGGAAACCTGTTATGATACAAATTCACAGAG 1963  
QY 1850 GTTGTGAGGATGATAGGAATGATGATGATGATGATGATGATGATGATGATGATGATG 1909  
Db 1964 ATTTGATGGCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2023  
QY 1910 TAACTTGAGGGGCTTGACGGCATTCAGGACCAAGTTTATGTTGGGAAGTGTGTTT 1969  
Db 2024 CAACATGAAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2083  
QY 1970 CAACAGAACAGCTATCTGTTATGATGATGATGATGATGATGATGATGATGATGATG 2029  
Db 2084 CAATAGGACAGCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2142  
QY 2030 GGCATCACTATGTTGGGGGCAAGAGGCAAGTCAAGGAAAGGAGGAGCTCAGATAA 2089  
Db 2143 ACATTTGTTTGAAGAGCTGCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2202

QY 2090 GAAAAAGTCGAACAAGCATGTC 2145  
Db 2203 GTCANAGCCGTATATGAAGAG 2262  
QY 2146 TAGAGGAGGTGTTGAAGGTG 2205  
Db 2263 TCAGGAGGCTATTTGAAGGT 2316  
QY 2206 TGAGCTTAGAGAAGAGATTG 2265  
Db 2317 GGAATTTGGAAGAAGCCTTGG 2376  
QY 2266 ATGCTGCTGTTCTCTCAGTCT 2325  
Db 2377 AAGGTGCATACCACTTCAAG 2436  
QY 2326 TAAAGTTGGCTATGAGGACAA 2385  
Db 2437 TCAGTTTGGATATGAGGACA 2496  
QY 2386 CTGTCACAGAAGATATTTTAA 2445  
Db 2497 CAGTTACGGAGGATATTTTGA 2556  
QY 2446 ATTGCATGCCCAAGCGCCAG 2505  
Db 2557 ACTGCATGCCACCAAGCCTT 2616  
QY 2506 TGAACCAAGTCTCGCGTGGG 2565  
Db 2617 TTAATCAGGTGCTCGTGGG 2676  
QY 2566 CCTATGCTATGGCTACGGAG 2625  
Db 2677 CTATCTGATGTTTACAATG 2736  
QY 2626 CCACCATTTCCCACTAACCT 2685  
Db 2737 CCATTTGTTTCCAATCACAT 2796  
QY 2686 GTCTGCTCAGTCGGAAGTTCA 2745  
Db 2797 GTCTTCTTACCAATAAATTA 2856  
QY 2746 TTGCGCTCTTCTTCAATTT 2805  
Db 2857 TTCTTCTTTCCGCTCCANN 2916  
QY 2806 GCATTCAGAGTGTGGAGGA 2865  
Db 2917 GCATTTGAAGATTTGGTGAGA 2976  
QY 2866 TGTTCGCTGCTTTTCAGGTC 2925  
Db 2977 TCTTCGAGGTTTCCAGGTC 3036  
QY 2926 TCACCTCAAAGGCTAATGACG 2985  
Db 3037 TTACCTCAAAGGCTATGATG 3096  
QY 2986 CGAGCTTCTCATCCCTCCGA 3045  
Db 3097 CCAGTTTGTCTATCTTCCGA 3156  
QY 3046 GCACCTCTAGCCCATCAACA 3105  
Db 3157 GAATTTGATGATGATGATG 3216  
QY 3106 TCTTTGCTTCTGGGTGATG 3165  
Db 3217 TCTTCTGATGCTGCTGNTCC 3276





1388 AAGAGCTTGGAGAAACGCTTTGGCCAGTCTCCAAATTTTATTGCATCCACCTTTATGACT 1447  
QY 2265 TATGGTGGTGTCTCAGTCCCTCCACTCCAGAAATCTCTTTGAAAGAGCTATCCATGTC 2324  
Db 1448 CAAGGTGGCATACCCCTTCAACAAACCCAGAGTTCCTGCTAAAGGAAGCTATACATGTC 1507  
QY 2325 ATAAGTTGTGGCTATGAGGACAAGTCTGAATGGGAACCTGAGATTGGTTGGATCTATGGA 2384  
Db 1508 ATTAGTTGGATATGAGATATAAACAAGATGGGGAAGAGATCGATGGATATATGCG 1567  
QY 2385 TCTGTACAGAAGATATCTTACTGAGTCAAGATCAGCAGAGAGCTGGCGTTTCAGTC 2444  
Db 1568 TCTGTACTGAAGATATTTAACTGTTTCAAGATCATGCAAGAGGTTGGATATCCATC 1627  
QY 2445 TATTGCATGCCAAGCGCCAGCTTTCAGGGATCTGCCCCCATCAATCTTTCAGATCGT 2504  
Db 1628 TACTGCATGCCACTTCCGGCTTCCTTCAAGGGTTCGCTCCGATTAATCTTCTGATCGT 1687  
QY 2505 CTGAACCAAGTGTGGGGTGGCTCTCGGTTCTGTTGAAATCTTTCAGCCGGCATTCG 2564  
Db 1688 CTCACCAAGTGTACGCTGGGCTCTTGGTTCAAGTTGAAATCTTACTTAGCAGACACTGT 1747  
QY 2565 CCCTTATGGTATGGCTACGAGAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAAC 2624  
Db 1748 CCTATCTGGTATGGTTACAAATGGAAGGCTAAAGCTTCTGGAGAGACTGGCATACATCAAC 1807  
QY 2625 ACCACCATTAACCACTTAACCTCTCCCGCTTCTAGTCTATGTTATGATGCTGCTATC 2684  
Db 1808 ACCATGTTTATCCAAATACATCTATCCCACTAGTAGTACTGCTCTTCTCGTATC 1867  
QY 2685 TGTCTGCTCACTGAAAGTTTCATCATCCAGAGATTAGCAACCTGGCCAGTATCTGTTTC 2744  
Db 1868 TGTCTACTCACCACAAATTTATTTCTCGGATTAGCAATTTGCTGGGGGCTTCTC 1927  
QY 2745 ATTGGCGTCTCTTCAATTTTCGCACTGGTATCTTGGAGATGAGTGGAGTGGTGT 2804  
Db 1928 ATCTGCTGTTTGTCTTCCATCTTCGCACTGTTTGGAGCTTCGATGGAGTGGTGT 1987  
QY 2805 GGCATTGACAGTGGTGGAGGATTAACAGTCTCTGGGTATCTGGAGGTATCTCTGCACAT 2864  
Db 1988 GGCATTGAGGATGGTGGAGAAATGAGCAGTTTGGGTATCTGGTGGCACCCTGCACAT 2047  
QY 2865 CTGTTTGGCGTCTTTCAGGCTCTTGAAGGTGCTTGGCGGTATCGCCGATATCGACACCACTTCAT 2924  
Db 2048 CTCCTTGTCTGTGTCCAAAGGTCTCTTAAAGTGTCTAGCAGGGATCGACACAACTTCACG 2107  
QY 2925 GTCACCTCAAGGCTAATGACGAAGAGCGACTTTTGTGAGCTCTACATGTTTCAAGTGG 2984  
Db 2108 GTCACATCAAGGCAACCGATGATGATGTTGTTGCTGAGCTGTATGTTTCAAGTGG 2167  
QY 2985 ACAGCGCTTCTATCCCTCCGAGCAATTTTGTATCAATTAACATGTTGGTGTGTTGCT 3044  
Db 2168 ACAACTCTTGTATCCCGCCACCACCTGCTGTGATTAACCTGTTGTTATAGTGGCT 2227  
QY 3045 GGCACCTCTACCCATCAACAGTGTACCAATATGCGGGCGCTCTTTGGAGAGCTC 3104  
Db 2228 GGAGTCTGTATGCTATCAACAGTGTCTACCAATATGCGGGTCCACTATTCGGGAAGCTG 2287  
QY 3105 TTCTTTGCTCTTGGGTGATGTTGCTACTTATACCAATCTCTCAAGGTCTTATGGGCAGG 3164  
Db 2288 TTCTTTGCAATCTGGGTGATCTCCACCTCTACCTTCTTCTGAGGGTCTCATGGGGAAG 2347  
QY 3165 CAAACCGCACACCGAGATGTCATGCTGGGCTGCTCTCTCGCTTCTATCTTCCTCC 3224  
Db 2348 CAGAACCGCACACCGACCATCTGTCATGTTTGGTCCGCTCTTCTGCTTCCATATCTCG 2407  
QY 3225 TTGCTGCGGTCTGCTTTCATTCATTCACCTACCCGT 3260  
Db 2408 CTGCTGTGGTGAAGATCGACCCCTTTCATATCCCT 2443

RESULT 6

AY103701  
LOCUS Zea mays PC0120363  
DEFINITION Zea mays  
ACCESSION AY103701  
VERSION AY103701.1 GI:21208  
KEYWORDS HTC.  
SOURCE Zea mays.  
ORGANISM Zea mays  
Eukaryota; Viridipl  
Spermatophyta; Magn  
clade; Panicoideae;  
1 (bases 1 to 3788)  
Hainey,C.F., Dolan,M  
Arthur,L.W., Hanafey  
Maize Mapping Projec  
Overgo Probes  
Unpublished (2002)  
REFERENCE 2 (bases 1 to 3788)  
AUTHORS Coe,E.C.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-20  
Missouri, Columbia, Mo  
FEATURES  
source  
1..3788  
/organism="Zea mays"  
/db\_xref="taxon:4577"  
/clone\_lib="PCC0120363"  
/clone\_lib="Library"  
/note="this assembly contigs to assembled mapping project  
BASE COUNT 851 a 954 c  
ORIGIN  
Query Match 36.4%;  
Best Local Similarity 65.3%;  
Matches 2129; Conservative

g 866 t 3 others  
are 1319.8; DB 11; Length 3788;  
ad. No. 0;  
Mismatches 1030; Indels 103; Gaps 9;  
CCAGATCTCGCGCGAGCGGCTGGGACACAGTT 148  
||||| ||| ||| ||| ||| ||| ||| ||| |||  
CCAGATATTCGGCGAGCGAGGTGCGGGTGGCTT 355  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
CGTCTCGCGCTTCCCGGTCTGCGCCCTGCTA 208  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
CGAGTGGCTTCCCGCTCTGCGCGCTTGGCTA 415  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
GGCTTCCCTCCAGTGCAGAACCAAGTACAAAGC 268  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
GGCGTCCCGCAGTGCAGGACCGCTTACAAGCG 475  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
GGAGGAGCGGAGGAGGAGGAGGAGGAGGAGG 328  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
GGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 535  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
TGGCCTGAGGAGGAGGAGGAGGAGGAGGAGG 385  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
CGGCCACGAGGAGGAGGAGGAGGAGGAGGAGG 595  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
CGGGGCGAGTGCAGTGTGGGCGACCCCAAGTA 445  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
CGGCCGCGGCGGCGGCGGCGGCGGCGGAGTCAA 655  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
GTATGACATGGAGAGATCCCTAGGGGATACGT 505  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
CAACG-----GCCAGATGGTTGATG 702  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
GAGAGAAATCCCTGGAGGCTTCGCTGATCATCA 565  
||| ||| ||| ||| ||| ||| ||| ||| ||| |||

8788 bp mRNA linear HTC 25-MAY-2002  
sequence.  
Streptophyta; Embryophyta; Tracheophyta;  
yta; Lillopsida; Poales; Poaceae; PACC  
pogoneae; Zea.  
ao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Morgante,M. and Tingey,S.V.  
Pont Consensus Sequences for Design of  
Maize Mapping Project, University of  
Missouri, Columbia, USA  
ers  
ays"  
B:637005"  
4577"  
3"  
Maize Mapping Project/DuPont Consensus  
ence is part of a project of EST  
ting from the application of public  
DuPont contigs; this resource was  
DuPont as part of a collaboration for the  
g of BACs in conjunction with the Maize  
g 866 t 3 others

Db 703 ACATCCGCGGAGCAGCAGCGCTCGTGCCGCTCTACATGGCGCGCGCGGCGCA 762  
QY 566 CATGATGTCCTCCCTACGGGAAACATCAGCAGAGTGCTCGTTCCCTATGTGAATCATTC 625  
Db 763 AGAGGATCCACCCGCTCCCTCGCTGATTCCAACCTTCCAGTGCAACCGAGATCCATGG 822  
QY 626 ACCAAATCCGTCGAAGGAGTTCTCCGGCAGTATTTGGGAATGTTGCCGTGGAAGAGAGAGT 685  
Db 823 ACCCGTCCAAGGA--TCTGGCGCCTACGGATATGCGAGCTGGCCTGGAAGAGAGAT 880  
QY 686 TGATGGTGGAAATGAAGCAGAGCAAGGGTGGGATTCCTCCATGACTAAATGGGACAAGCAT 745  
Db 881 GGAGGGTGGAGCAGAGCAGGAGCGCCTGCGAGATGTCAGGAGCGAGG----- 931  
QY 746 TGCTCCCTCTGAAGGTCGGCAGCTACTGACATCGATGATCTACTGAATACAACATGGA 805  
Db 932 -----TGCGGTGATTTGGGATGGCGAGCATGC 958  
QY 806 AGACGCTTTTACTGAATGATGAACCTCGCCAGCCTCTATCTAGAAAAGTCCCATTTGCTTC 865  
Db 959 AGATCTGCCACTAATGGATGAAGCTAGGCAGCAATTTCCAGAAAAGTCCCTATATCATC 1018  
QY 866 CTCCAAATAATCCCTACAGAATGTCATGTTCTGCGGTGGTGGTTTAAAGCATCTT 925  
Db 1019 AAGCGAATTAATCCCTACAGGATGATATCGTTATCCGTTGGTGGTTTGGGTTCTT 1078  
QY 926 CTTGCACTACCGTCTCACAAATCCTGTGCGTAATGCATACCCACTGTGCTTTATCTGT 985  
Db 1079 CTTCCACTACCGAGTATGATCCCGGAAAGATGATTTGCATTTGGCTCATATCTGT 1138  
QY 986 TATATGTGAGATTTGTTTGTCTTATCTGGATCTGATCTGATCTCCGAGTGGTTTCC 1045  
Db 1139 AATCTGTGAATCTGTTGGGATGTCCTGGATCTTGATCAGTTCCCAAGTGGCTTCC 1198  
QY 1046 AATCAACCGGGAGACCTACCTTGATAGATGGCTTTAAGTATGACGAGAGGTGAACC 1105  
Db 1199 AATCGAGAGAGACTTACCTGGACGTTTGTCACTAAGGTTTGACAAGAGGTCAACC 1258  
QY 1106 GTCTCAGTTGGCTGCTGTGACATATTTGTCAGTACAGTCGACCCCTTGAAGAGCCACC 1165  
Db 1259 CTTCTAGTTGCTTCCCAATCGACTTCTGTGACAGGTTGATPCCCAAGAGAACTCC 1318  
QY 1166 TATCGTCACTGCCAACACTGTGCTATCCATTTCTGCTGATTTATCCCGTGACAAAGT 1225  
Db 1319 CTTGGTCACAGCAACACTGCTCTTCCATCTTCTGTGGATATCCCGTTGAGAGGT 1378  
QY 1226 CTTCTGCTATGATATGATGACGAGCTTCAATGCTGACTTTTGACGCAATGGCTGAGAC 1285  
Db 1379 CTTCTGCTATGTTCTGATGATGGTGTGCAATGCTTACGTTTGAAGCATTTGTCTGAAC 1438  
QY 1286 TTCAGAGTTTGTAGGAATGGGTACCAATTTGTGAAGAGATGATGACATTTGAACCCAGAC 1345  
Db 1439 ATCTGAATTTGCAAGAAATGGGTTCTCTTTCAGCAAAAAGTTTAATATCAGCCTCGTC 1498  
QY 1346 TCCCGAGTTTACTTTTGCCAGAAAATTGATTTACCTTGAAGACAAAAGTCAGCCTTCATT 1405  
Db 1499 TCTGAGTGTACTTCCACAGAAAGATAGACTACCTTGAAGACAAAGGTTGCTGCTTCATT 1558  
QY 1406 TGTAAAGACCGCGGCCCATGAAGAGAAATATGAAGAAATTTAAATCAGGATAAATGC 1465  
Db 1559 TGTAGGAGAGAGGCGCATGAAGAGAAATACGAGGAATTCAGGTAAGGATCAATGC 1618  
QY 1466 CTTAGTTTCTAAGGCATTGAAATGCCGAGGAAGATGGATCATCGAATGGCACACC 1525  
Db 1619 CTTGGTTGCAAAAAGCCCAAGGTTCTCTGAGGAAGATGGACAATCAAGATGGAAGCCC 1678  
QY 1526 ATGCCAGGAAACAATACCAAGGATCATCTGGAATGATTCAGTTTTCCTTGGTCACAG 1585  
Db 1679 CTGGCTTGGAAACAACGTAACCGCATCATCTGGAATGATTCAGGTAATCTTGGCCAAAG 1738  
QY 1586 TGTGGCCTTGTACTGAGGGTAATGAGCTCCCGCTTTAGTTTATGTCTGCTGTAATA 1645  
Db 1739 TGGCGGTCGTGATGTGAAGGAATCAGTTGCTCGCCTCGCTTATGCTCGAGAGAAA 1798

QY 1646 GCGTCTCGGTTCCAGCACCAAGAGGCTGGTCCATGAATGCCCTTGTTCGTCTC 1705  
Db 1799 GAGGCCAGGTTATTAACCATCAAGAAGGCTGGTCCATGAATGCATGGTCCGTGCTC 1858  
QY 1706 AGCTGCTCTTACTAATGGACAATACATGTTGAATCTTGAATTTGTGATCACTACATCAAAA 1765  
Db 1859 TGTCTCTTATCAAAATGCTGCATACCTATTGAACCTGGACTGTGATCACTACATCAACAA 1918  
QY 1766 CAGCAAGGCTGTCCGAGAAGCTATGTGCTTCTTAATGGATCCAAACCTAGTCCGCAAGT 1825  
Db 1919 TAGCAAGGCCATAAAAGAGGCTATGTGTTTCATGATGGATCTTTGGTGGGAAGAAAGT 1978  
QY 1826 CTGTTATGTGAGTTCCCAACAAAGGTTGATGGGATTCATAGGAATGATCATATGCAAA 1885  
Db 1979 GTCTATGCTACAGTCCCTCAGAGTTGATGTTGATGACAAAATGATCATAGCTAA 2038  
QY 1886 CAGGAACACTGTCTTTTTTGTATTAATCTTGAGGGCCTTGACGGCATTCAGGACCAAGT 1945  
Db 2039 CAGGAACCTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGTTATTCAGGACCAT 2098  
QY 1946 TTATGTGGGAAGTGGTTGTTTTCACAGAACAGCTATCTATGTTATGAGCCCCCAAT 2005  
Db 2099 TTATGTGGTACTGGATGTTTTCAGACGGCAGGCACTGTATGTTATGATGCTCTTAA 2158  
QY 2006 TAAGCGGAAGAAGCCA-----GGTTTCTTTGGCATCACT 2038  
Db 2159 ACGAAGAACGCCACCTCAAGAACTTGCAACTGTGGCCCAAGTGGTGCCTCTCTTGTGTG 2218  
QY 2039 ATGTGGGGCAGAAAGAGGCAAGCAAGTCAAAGAAAAGAGAGCTCAGATAAGAAAAA--- 2095  
Db 2219 CTCGACGAGGAACAAGATAAAGAAAGACACTACAAAACAAAGAGGAGAGAAAG 2278  
QY 2096 ---GTCGAACAAGCATGTGGACAGTTCTGTTCCAGTATTCATCTCGAAGACATAGAGGA 2152  
Db 2279 ATTATTTTCAAGAAAGCAGAAACCCATCTCTGCATATGCTTTGGGTGAATTTGATGA 2338  
QY 2153 GGGTCTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATGCTCAAAATGAGCTT 2212  
Db 2339 AGGTCTCCAGGTGC-----TGATATCGAAGGCCGGAATCGTAATCAACAGAACT 2392  
QY 2213 AGAAGAGAGATTTGGCCAGTCAGCAGCAATTTTGGCTTCCACTCTGTATGGAAATATGGTG 2272  
Db 2393 AGAAGAGAAATTTGGGCAAGTCTCTGTTTGTGTCATCAACACTTCTTGAGAAGGAGG 2452  
QY 2273 TGTTCCTCAGTCTCCATCCAGAAATCTCTTTTGAAGAAGCATATCCATATCAATAGTTG 2332  
Db 2453 GACCCCTGAAGAGCGCAAGTCCAGCTTCTTCTTGAAGGAAGCATATACATGTTATCAGCTG 2512  
QY 2333 TGGCTATGAGGACAAGTCTGAATGGGAACTCAGATTTGGTGGATCTATGATCTGTAC 2392  
Db 2513 CGGCTACGAAGACAAGACCGACTGGGGAAGAGAGATTTGGTGGATTTACGATCGATCAC 2572  
QY 2393 AGAAGATATTTACTTGGATTCAAGATGCAGCAAGAGGCTGGGTTTCAGTCTATTGCAAT 2452  
Db 2573 AGAGATATCTGACTGATTTAAGATGCATGCCATGGCTGGCGGTCTATTTACTGAT 2632  
QY 2453 GCCCAAGGCCAGCTTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGATCTCTGGAACA 2512  
Db 2633 CCCGAAGCGGCTGCTATTCAAAGGTTCTGCGCTCTGAAACCTTTCGACCCGCTCTCACCA 2692  
QY 2513 AGTGTCTGGTGGGCTCTCGGTTCTGTTGAAATCT--TTTCAGCCGGCATTTGCCCTTAT 2571  
Db 2693 GGTCTTCTGGTGGGCTTGGGTCGGTGGAAATTTTCTTCAGNANCAATGGNCACTTT 2752  
QY 2572 GGTATGGCTACGAGGCGGCTCAAGTTCTCTGGAGAGATTCGCTTACATCAACACACCA 2631  
Db 2753 GGTACGGATACGCGCGGCTTAAATTCCTGGAAAGGTTTCTTATATCAACTCCATCG 2812  
QY 2632 TTTTACCCACTAACCTCTCTCCCGCTTCTAGTCTATTTGTTATATGCGCTGCTATCTGTCTGC 2691  
Db 2813 TTTATCCCTGGAGCTCCATTCCTCTCTCGGCTTACTGTACCTTGCTGCCATCTGCCCTGC 2872

QY 2692 TCACTGGAAGTTTCATCATCCAGAGATTAGCAACTTGGCCAGTATCTGTTTATTGCGC 2751  
Db 2873 TCACGGGAAGTTTATCACACAGGAGCTTACCAATTCGCCAGTATCTGTTTATGSCAC 2932  
QY 2752 TCCTTCCTTTCAATTTTCGCCACATGGTATCTTGCAGATGAGGTGGAGTGGTGGCATG 2811  
Db 2933 TTTTCTATCTCATCTCCGTGACCGCATCTCTGAAATGAGTGGAGTGGCGTGGCCATCG 2992  
QY 2812 ACAGTGGTGGAGGATGAACAGTTCTTGGGTCAATTTGGAGTATCTCTGCACATCTGTTTG 2871  
Db 2993 ACAGTGGTGGAGGAGGAGGAGTCTTGGGTCAATTCGGAGGCGTTTCGGCCATCTGTTG 3052  
QY 2872 CCCTCTTTTCAAGGTCTTCTGAAGGTGCTTGGCCGTATCGACACCAACTTCACTGTACCT 2931  
Db 3053 CGGTGTTTCAGGGCCCTGCTGAAGGTGTTTCGGCCATCTGACAGCTTCAACGCTGACGT 3112  
QY 2932 CAAAGGCTAATGACGAGGAGGAGTCTTGTCTGAGCTCTACATGTTCAAGTGGACGACGC 2991  
Db 3113 CGAAGGCGGGGACGAGCA---GGAGTTCTGGAGCTGTACAGCTTCAAGTGGACCAACC 3169  
QY 2992 TTCTCATCCCTCCGACGACCAATTTTATCATTAACATGTTGGTGTGCTTGTCTGGCACCT 3051  
Db 3170 TGCTGATACCCCGGACACGCTCTCTCTGCTGAACTTTCATCGGGTGGTGGCGGGATCT 3229  
QY 3052 CCTACGCCATACAGAGTGTACCAATCATGGGGCGGCTCTTTGGGAAGCTCTTCTTTG 3111  
Db 3230 CGAAGCGATCAACAACAGGAGTACGAGTCTGTTGGGGCCCTGTTTCGGGAAGCTTCTTCTG 3289  
QY 3112 CCTTCTGGGTGATGTTCACTTATACCAATCTCTCAAGGTCTTATGGGAGGCAAAACC 3171  
Db 3290 CGTCTGGGTGATCTCCACCTGATCCCGTCTCTCAAGGTCTTGGTGGGAGGAGCAACA 3349  
QY 3172 GCACACGAGATGTCATGCTGTGGGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3231  
Db 3350 GGACGCGGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3409  
QY 3232 GGGTTCGTGTGATTCATTCATCTACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3291  
Db 3410 GGGTCCGCGTCGACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3469  
QY 3292 TCAACTGCTAGGAAAGTGGAG 3313  
Db 3470 TGGACTGCAACTGAAGTGGGG 3491

RESULT 7  
AY104730 3783 bp mRNA linear HTC 25-MAY-2002  
LOCUS  
DEFINITION Zea mays PCO100501 mRNA sequence.  
ACCESSION AY104730  
VERSION AY104730.1 GI:21207808  
KEYWORDS HTC.  
SOURCE Zea mays.  
ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC  
Clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 3783)  
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
Maize Mapping Project/DuPont Consensus Sequences for Design of  
Overgo Probes

Unpublished (2002)  
2 (bases 1 to 3783)  
Coe,E.C.  
Direct Submission  
Submitted (25-APR-2002) Maize Mapping Project, University of  
Missouri, Columbia, MO 65211, USA

FEATURES  
Location/Qualifiers  
1..3783  
/organism="Zea mays"  
/db\_xref="MaizeDB:635765"  
/db\_xref="taxon:4577"

BASE COUNT 933 a 908 c 2 g 940 t  
ORIGIN  
Query Match 36.3%  
Best Local Similarity 64.4%  
Matches 2132; Conservative  
QY 47 GCTCTGCCGCGCCATGACACGG 363  
Db 233 GCTCGTGTGTCATCGCGCGGAT 363  
QY 107 GGACGTGTGCCAGATCTGCGCG 363  
Db 293 CCAGGTGTGCCAGATTTGCGGG 363  
QY 167 CGCTTGGACGTCTGCGCGTTG 363  
Db 353 GGCTTGCACGAGTTCGCGCTTC 363  
QY 227 GGCACCCAGCGCTGCTCCAC 363  
Db 413 GGGCAGCGAGTTCGCGCGCG 363  
QY 287 GATTCGCGGGGAGGAGCGCG 363  
Db 473 CGTCCCGGGGAGGAGGAGG 363  
QY 347 TGATCTGGACATGAGGAGGAGG 363  
Db 528 ----TGGAGCGACAAGCACGAG 363  
QY 407 GAACACCGGGGCGAGTGGCAAT 363  
Db 584 GAGCTAGGCGCGCGCGCGCGAG 363  
QY 467 CTCGAAGTATGACAGTGGAGG 363  
Db 644 TGTTCCTCTCTCACCAACGGA 363  
QY 527 GATGTAGGAGAAATCCCTGGA 363  
Db 704 TGTGCGCTCTGTCGTGGTGGG 363  
QY 587 CATCAGCAGACGTGCTCCGTTT 363  
Db 761 TCCCAACCTCTCTGTGCAACCG 363  
QY 647 CTCGCGCAGTATTTGGGAATGT 363  
Db 821 CTAC-----GGGAGCGTA 363  
QY 707 GGACAGGGTGGCATTTCCCATG 363  
Db 872 GGAGAGGATGCACACGACGAG 363  
QY 767 AGCTACTGACATCGATGATCT 363  
Db 923 -----TGCAGATCTACCACTATGATGA 363  
QY 827 AACTCGCCAGCCTCTATCTAGA 363  
Db 947 AGCTAGACAGCCATTTGTCCAG 363  
QY 887 ATGGGTATTTGTTCTCGGGTTG 363

01"  
ze Mapping Project/DuPont Consensus  
ence is part of a project of EST  
ating from the application of public  
DuPont contigs; this resource was  
ont as part of a collaboration for the  
g of BACs in conjunction with the Maize  
2 g 940 t

ore 1315.6; DB 11; Length 3783;  
ed. No. 0;  
Mismatches 1074; Indels 102; Gaps 7;  
GGAGGCCCTGAAGTCCGGGAGGACGCGGGCGG 106  
AGCCAGGGGCCAAGCCCATGGACCAGCGGACGG 292  
SCCTGGGCACACGTTGGACGCGGACGCTTTCAC 166  
ACGTGGGGCGCAACCCCGACGCGGGAGCGTTCGT 352  
CTCGCGCCCTGTCTACGACGACGAGCGCAAGGA 226  
CTCGCGGAGCTGTCTACGATACGAGCGCGCGA 412  
AGACAAGTACAAGCGGCCACAGAGGAGGAGCCAGC 286  
AGACCCGCTTCAAGCGGCTCAAGGGGTGCGCGCG 472  
CTGATGCGGATGATGGTGTAGTGTCAACTACCC 346  
SCGTGACGACCTGGAGAACGAGTTCAC----- 527  
AGAAGATTGTGACAGGATGCGAGTGGGCGCAT 406  
AGTACCTCGCGAGTTCATGTCACGCGCCACAT 583  
CCACCCCAAGTATGACAGTGGCGAGATCGGCGCT 466  
ACGGGTGCGCGACGCCATTCACCCCATCCCAA 643  
TAGGGGATAGTCCCTTCTAGTCAACCAAGCCCA 526  
TGGTGTGATGACATCCCGCGGACGACGCGCT 703  
CGCTGTATCATCATGATGTCCTCTAGCGGGGAA 586  
GGAAGAGGATTCACCTCT---CCGTACGCGGA 760  
ATGTGAATCATTCACCAAAATCCGTCAAGGGAGTT 646  
CTATGACCCCTTCCAGGATCTCGCGCATATGG 820  
GGAAGAGAGAGTTTGGTGGTGAATAATGAAGCA 706  
GGAAGAGAGAGTGGAGAGCTGGAAGCAGAAGCA 871  
ATGGACAAGCATTTGCTCCCTCTGAAGTGGCGG 766  
ATGGCGGCGGAGTATGATGATGA----- 922  
AATACAACATGGAAGACGCTTTTACTGAATGATGA 826  
-----TGCAGATCTACCACTATGATGA 946  
CCCCCATTTGCTTCTCCAAATAATCCCTACAG 886  
CCCGCTTCTTCAAGCAAAATCAACCCCTATAG 1006  
CTTAAGCATCTTCTCGACTACGCTCTCAAA 946



Qy	3074	CCAATCATGGGGGGCGCTCTTTGGGAAGCTCTTCTTTGCCCTCTGGGTGATGTTCAC	3133
Db	3224	TGAATCATGGGGGGCGCTTTTCGGGAAGCTCTTCTTTGCA	3283
Qy	3134	ATACCCATTCTCAAGGCTCTATGGGCAGCGAACAACCGCACACCGAGATTGTCATCGT	3193
Db	3284	TTACCCGTTCTCTCAAGGCTCTGTTGGGAGGCGAACAAGACGCAACGATGTCATTTGT	3343
Qy	3194	CTGGGCTCTCCTCCCTCGCTTCTATCTTCTCTGCTGTGGGTTCGTGTGATCCATCCAC	3253
Db	3344	CTGTGTCCATCTCTCTGGCTTCGATCTTCTCGCTGCTTTTGGGTCCGGATCGACCGCTTCT	3403
Qy	3254	TACCCGCTCTCGTGGCCCAATATCCAAACCTCTGGGCATCAACTGCTAGGAAAGTGGAG	3313
Db	3404	TGCGAAGGATGATGTCCTCCCTGTTGGAGGAGTGTGGTGTGATTGCAACTAGGAGGTGAG	3463
Qy	3314	TTTGTAGA	3321
Db	3464	CACGTGGA	3471
RESULT	8		
AY103655			
LOCUS	AY103655	3897 bp	mrna linear HTC 25-MAY-2002
DEFINITION	Zea mays PCO096398 mRNA sequence.		
ACCESSION	AY103655		
VERSION	AY103655.1 GI:21206733		
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 3897) Hainey,C.F., Dolan,M., Mao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.		
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes		
JOURNAL	Unpublished (2002)		
REFERENCE	2 (bases 1 to 3897) Coe,E.C.		
AUTHORS	Direct Submission		
TITLE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
JOURNAL	Missouri, Columbia, MO 65211, USA		
FEATURES	Location/Qualifiers		
source	1..3897		
	/organism="Zea mays"		
	/db_xref="MaizeDB:635496"		
	/db_xref="taxon.4577"		
	/clone="PCO096398"		
	/clone_lib="Maize Mapping Project/DuPont Consensus Library"		
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
BASE COUNT	982 a	839 c	1008 g
ORIGIN	1068 t		
Query Match	35.0%;	Score 1267.8;	DB 11; Length 3897;
Best Local Similarity	63.9%;	Prod. No. 0;	
Matches 2108; Conservative	0;	Mismatches 1082;	Indels 111; Gaps 8;
Qy	47	GCTCTCGCGCGCATGACGCGACGCGCCCTCAAGTCTCGGGAGGACGCGGCGG	106
Db	203	GCTCGTCTATCGCGCGCGACGGCGATCCGGGCCGAAGCCCGCGGAGCAGAACGG	262
Qy	107	GGAGTGTGCAGATCTGCCCGCACGCGCTTGGGACCACTGTGGACGCGACGCTTTTAC	166
Db	263	GCAGTGTGCAGATTTGCGCGACGACGCTTGGCCCGCGGCGGCGGACCCCTTCTGT	322



Qy	1247	CGAGACTTCAATGCTGACCTTTTGTAGCGCAATTGGCTGAGACTTCAGAGTTTGTCTAGGAATG	1306
Db	1334	TGTTGCTGCAATGCTAACCTTTTGAAGCAATTAATCTGAAACATCTGAAATTTGCAAAAGAAATG	1393
Qy	1307	GGTACCATTGTGGAAGAAGTATGACATTGAACCCAGAGCTCCCGAGTTTGTACTTTTCCCA	1366
Db	1394	GGTTCCTTTCTGCAAAAGCGTGCAATATATGAACCTCGCGCTCCAGAGTGGTACTTTCCAACA	1453
Qy	1367	GAAATTTGATTTACCTGAAAGACAAAAGTCCAGGCTTTCAATTTGTATAAGAGCCGCCGGCCAT	1426
Db	1454	GAAGATAGACTACTTGAAGACAAAGTGCCACGCAAACTTTGTTAGGAGAGGAGAGCAAT	1513
Qy	1427	GAAGAGAGAAATATGAAGAAATTTAAATCAGGATAAATGCCCTAGTTTCTTAAGGCATTGAA	1486
Db	1514	GAAGAGAGAGTATGAGGAATTCAAAGTGAAGATCAATGCCTTAGTTGCGCAAGGCCAGAA	1573
Qy	1487	AGTCCCGGAGGAGGATGATCATCGAGTGGCACACCATGGCCAGGAACAATACCAG	1546
Db	1574	AGTTCTGTGAAGAAGATGAGCAATCGAGATGGAACCCCTGGCCCTGGAACAATGTTTCG	1633
Qy	1547	GGATCATCTCTGGAATGATTCAGGTTTTCTTGGTCACAGTGGTGGCCTTGATCTAGAGG	1606
Db	1634	TGATCATCTCTGGAATGATTCAGGTTCTCTTGGCCAAAGCGGAGCCTTGACTGTAGAGG	1693
Qy	1607	TAATGAGCTCCCCGTTTATGTTATGTGTCGTGAAAGCGTCTCGGTTTCCAGCACCA	1666
Db	1694	AAATGAACTGCCACGATTGGTTTATGTTTCTAGAGAGAAACGACGAGGCTATAACCATCA	1753
Qy	1667	CAAGAAGGCTGGGCCATGAATGCCCTTGTTCGTCTCAGCTGTCCTTACTAATGACA	1726
Db	1754	TAAGAAGCTGGTGTATGAATGCAATGGTCGAGTCTCTGCTGTACTTAACAAATGCTCC	1813
Qy	1727	ATACATGTTGAATCTTGTATGTTGATCACTACATCAACACAGCAAGGCTGTCGGAAGC	1786
Db	1814	ATATTGTTAACTTGGATTTGATCACTACATCAACACAGCAAGGCTATAAGAGAGC	1873
Qy	1787	TATGTGCTTCCTAATGGATCCAAACCTAGTCCGCAAGTCTGTATGTCAGTTTCCACCA	1846
Db	1874	AATGTGTTTATGATGGACCCCTTACTAGGAAGAAGGTTTGTCTATGTACAGTTTCCCTCA	1933
Qy	1847	AAGTTTGTATGGGATGATAGGAATGATGATATCAACACAGGACACTGCTCTTTTTTGA	1906
Db	1934	AGATTGTATGGGATTTGATCGCCATGACCGATATGCTAACCGGAATGTTGTCTTTTTTGA	1993
Qy	1907	TATTAACCTTGAGGGCCTTGACGGCATTCAGAGCACGTTTATGTGGAACTGCTGTGTGT	1966
Db	1994	TATCAACATGAAGGTTTGGATGGTATTCAGGTTCCAATTTATGTTGCTGCTGATGTGT	2053
Qy	1967	TTTCAACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAGCGGAAGAAGCCA-----	2021
Db	2054	ATTTAGAAGCGAGGCATTAATATGGTTATGATGCCCCCAAAACAAGAAGACCCACCATCAAG	2113
Qy	2022	-----GGTTCTTGTCATCACTATGTGGGGCGCAAGAGAGGC	2059
Db	2114	GACTTGCACTGCTGGCCCAAGTGGTGTCTTTGTGCTGTTTGGCAATAGGAAGCA	2173
Qy	2060	AAGCAAGTCAAAGAAAAGGAGCTCAGATAAGAAAAAG-----TCGAACAAGCATGTGGA	2113
Db	2174	AAGAAGACTACCAACCCCAACACAGAGAAGAAAAGTATTATTTTTCAGAAGAAGAGA	2233
Qy	2114	CAGTTCTGTTCCAGTATTCAATCTCGAAGACATAGAGGAGGTTGTGAAGTGTCTGGGTT	2173
Db	2234	GAACCAATCCCTTGCAATATGCTCTTGGTGAATTTGACGAAGCTGCTCCAGGAGC-----	2287
Qy	2174	TGATGATGAGAAATCAGTTTCTCATGCTCAATGACCTTAGAGAAGAGATTTGGCCAGTC	2233
Db	2288	TGAGAATGAAAGGCCGATTTGTAAATCAACAAAAATTAGAAAAGAAATTTGGCCCAATC	2347
Qy	2234	AGCAGATTTTGTGGCTCCACTCTGATGGAATATGGTGGTGTCTCTCAGTCTCCACTCC	2293
Db	2348	TTCTGTTTGTGTACATCCACACTCTCTCGAAGATGCTGGAACCTTGAAGAGTGCAGTGCC	2407
Qy	2294	AGAATCTCTTTTGAAGAAGAGCTATCCATGCTCAATAAGTTGTGGCTATGAGGACAAAGTCTGA	2353

[illegible]

[illegible]

RESULT 11					
BQ802778					
LOCUS		762 bp	mRNA	linear	EST 30-JUL-2002
DEFINITION	WHE2829_H06_O11S Triticum monococcum vernalized apex cDNA library				
	Triticum monococcum cDNA clone WHE2829_H06_O11, mRNA sequence.				
ACCESSION	BQ802778				

Qy	805	AAGAGCGCTTTACTGAATGATGAAACTGCCAGCCCTCTATCTAGAAAAAGTCCCCATTCCT	804
Db	191	AAGAGCGCTTTACTGAATGATGAAACTGCCAGCCCTCTATCTAGAAAAAGTCCCCATTCCT	250
Qy	865	CCTCCAAATTAATGCCCTACAGATGGTCATTTGGTCTCGGGTGGTGTCTCTAAGCATCT	924
Db	251	CCTCCAAATTAATGCCCTACAGATGGTCATTTGGTCTCGGGTGGTGTCTCTAAGCATCT	310

```
QY 925 TCCTGCACTACCGTCTCACAAATCCCTGCTGCTAATGCATACCCACCTGTGGCTTTTATCTG 984
|||||
Db 311 TCCTGCACTACCGTCTCACAAATCCCTGCTGCTAATGCATACCCACCTGTGGCTTTTATCTG 370
QY 985 TTATATGTGAGATTGGTTGCTTTTATCTGCTGATGATGATGATGATGATGATGATGATGATG 1044
|||||
Db 371 TTATATGTGAGATTGGTTGCTTTTATCTGCTGATGATGATGATGATGATGATGATGATGATG 430
QY 1045 CAATCAACCGGGAGACCTACCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1104
|||||
Db 431 CAATCAACCGGGAGACCTACCTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 490
QY 1105 CGTCTCAGTTGGCTGCTGTGTGACATATTTGTCTAGTACAGTCGACCCCTTGAAGGAGCCAC 1164
|||||
Db 491 CGTCTCAGTTGGCTGCTGTGTGACATATTTGTCTAGTACAGTCGACCCCTTGAAGGAGCCAC 550
QY 1165 CTATCTCACTGCAACACTGTCTATCCATTTCTGCTGTGATTTATCCCGTGGACAAGG 1224
|||||
Db 551 CTATCTCACTGCAACACTGTCTATCCATTTCTGCTGTGATTTATCCCGTGGACAAGG 610
QY 1225 TCTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284
|||||
Db 611 TCTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 670
QY 1285 CTTCAAGTTTGCTAGGAATGGTACCATTGTTGGAAGAGTATGATGATGATGATGATGATGATG 1344
|||||
Db 671 CTTCAAGTTTGCTAGGAATGGTACCATTGTTGGAAGAGTATGATGATGATGATGATGATGATG 730
QY 1345 CTCCCGAGTTTACTTTTCCAGAAAATTGAT 1376
|||||
Db 731 CTCCCGAGTTTACTTTTCCAGAAAATTGAT 762

RESULT 12
BG368813
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
Source
1..870
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEi0020M13f"
/hvcdna_lib="Hordeum vulgare 20 DAP spike EST library
HVSMEi0020M13f"
/tissue_type="20 DAP spike"

/lab_host="
/notes="vector
plants were
California
spikes with
). Total RNA
primary unan
phu were in
phagemids f
California
picked at t
(Begum, Pal
preparation
performed a
, Rambo, Ma
phred sequ
phred value
http://www
this clone
see close
Genetically
Barley gene
(http://whe
BASE COUNT 230 a 174 c
ORIGIN
Query Match 20.1%
Best Local Similarity 93.8%
Matches 813; Conservative
QY 1747 GTGATCACTACATCAACAACAG 1806
Db 1 GTGATCACTACATCAACAACAG 60
QY 1807 CAAACCTAGTCCGCAAGTCTG 1866
Db 61 CAAACCTAGTCCGCAAGTCTG 120
QY 1867 GGAATGATCGATATGCAACAG 1926
Db 121 GGAATGATCGATATGCAACAG 180
QY 1927 ACGGATTCGAAGCAGCAGTTT 1986
Db 181 ATGGCATTCAAGGCCAGTTT 240
QY 1987 ATGGTATGAGCCGCCAATTAT 2046
Db 241 ATGGTATGAGCCGCCGATTAT 300
QY 2047 GCAAGAAGAGGCAAGCAAGTC 2106
Db 301 GCAAGAAGAGGCAAGCAAGTC 360
QY 2107 ATGTGACAGTCTCTGTTCAGT 2166
Db 361 ATGTGACAGTCTCTGTTCAGT 420
QY 2167 CTGGGTTTGATGATGAGAAATG 2226
Db 421 CTGGGTTTGATGATGAGAAATG 480
QY 2227 GCAGTCAGCAGCATTTGTTGG 2286
Db 481 GCAGTCAGCAGCATTTGTTGG 540
QY 2287 CCACTCCAGAAATCTCTTTTGA 2346
Db 541 CCACTCCAGAAATCTCTTTTGA 599
QY 2347 AGTCTGAATGGGAACTGAGAT 2406
|||||
```

```
LambdaZAP; Site 1: EcoRI; Site 2: XhoI;
in the greenhouse at the University of
arside (Fenton, SJ Close, TJ Close). Whole
s trimmed were collected at 20 DAP (Fenton
s prepared, poly(A) RNA was purified, one
ied cDNA library was made, and 1 million
s excised to give pBluescript SK(-) cDNA
s TJ Close lab at the University of
arside (Choi). Phagemids were plated and
lemson University Genomics Institute (CUOI)
Frisch, Atkins and Wing). Plasmid DNA
NA sequencing and sequence analysis were
SI (Wing, Yu, Frisch, Henry, Simmons, Oates
The sequence has been trimmed to remove
and contains a minimum of 100 bases of
or above. For more details on library
sequence analysis see
ne.Clemson.edu/projects/barley. To order
http://www.genome.clemson.edu/orders Also
ing R, Kleinhofs A, Wise R (2001)
Physically anchored EST resources for
Barley Genetics Newsletter 31:29-30.
y.usda.gov/ggpages/bgn/31/cover.html."
s g 238 t 3 others
ore 729.4; DB 12; Length 870;
ad. No. 1.7e-192;
Mismatches 48; Indels 6; Gaps 5;
SCTGTCGGAAGCTATGCTCTCTCTATGATC 1806
|||||
SCTGTCGGAAGCTATGCTCTCTCTATGATC 60
STGCAGTTCACAAAGTTTGATGGATGATA 1866
|||||
STGCAGTTCACAAAGTTTGATGGATGATA 120
ACTGCTTTTTTGATATTAACCTGAGGGCCCTG 1926
|||||
ACTGCTTTTTTGATATTAACCTGAGGGCCCTG 180
SGAACTGTTGTTTTTCAACAGAACGCTATCT 1986
|||||
SGAACTGTTGTTTTTCAACAGAACGCTATCT 240
NAGAAGCCAGTTTCTTGGCATCATCTGCTGGG 2046
|||||
NAGAAGCCAGTTTCTTGGCATCATCTGCTGGG 300
AAAAGAGCTCAGATAAGAAAAGTTCGAACAAGC 2106
|||||
AAAAGAGCTCAGATAAGAAAAGTTCGAACAAGC 360
AATCTCGAAGACATAGAGGAGGGTGTGAAGGTG 2166
|||||
AATCTCGAAGACATAGAGGAGGGTGTGAAGGTG 420
TTCATGCTCAATGAGCTTAGAGAGAGATTTG 2226
|||||
TTCATGCTCAATGAGCTTAGAGAGAGATTTG 480
ACTCTGATGAATATGTTGGTGTCTCTCAGTCCT 2286
|||||
ACTCTGATGAATATGTTGGTGTCTCTCAGTCCT 540
SCTATCCATGTCATAAGTTGTGGTGTATGAGGACA 2346
|||||
SCTATCCATGTCATAAGTTGTGGTGTATGAGGACA 599
TGGATCTATGATCTGTACAGAGAGATTTCTTA 2406
|||||
```



**AUTHORS** Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Nguyen, H.T., Raush, C.J., Seaton, C.L., Tong, J.C., and Zhang, D.

**TITLE** The structure and function of the expressed portion of the wheat genomes - Normalized shoot cDNA library

**JOURNAL COMMENT** Unpublished (2000)  
Contact: Olin Anderson  
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center  
800 Buchanan Street, Albany, CA 94710, USA  
Tel: 5105595773  
Fax: 5105595818  
Email: oanderson@pw.usda.gov  
Sequences have been trimmed to remove vector sequence and low quality sequence with phred score less than 20  
Seq primer: SK primer

**FEATURES**

source	Location/Qualifiers
1..710	/organism="Triticum aestivum"
	/cultivar="Chinese Spring"
	/db_xref="taxon:4565"
	/clone="WHE3028_D08_H16"
	/clone_lib="wheat unstressed seedling shoot normalized cDNA library"
	/tissue_type="Etiolated shoot"
	/dev_stage="Five day old seedling"
	/lab_host="E. coli DH10B"
	/note="Vector: Lambda Uni-ZAP XR, excised phagemid pBluescript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized, germinated and grown aseptically in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Roots were harvested. The tissue, total RNA, and poly(A) RNA were prepared. A cDNA library was made in the pTZ Close lab (Choi, Close, Fenton) at the University of California, Riverside. The cDNA clones were in vivo excised to give pBluescript phagemids before normalization was carried out. The mass excision of phagemid library and normalization were done in HT Nguyen lab by D. Zhang at Texas Tech University. Normalization protocol used was that of Soares'. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
BASE COUNT	139 a 189 c 166 g 216 t
ORIGIN	
Query Match	19.0%; Score 689.4; DB 14; Length 710;
Best Local Similarity	99.1%; Pred. No. 2.5e-181;
Matches	693; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 2578	GCTACGAGGCGGCTCAAGTTCCTGGAGAGATTGGCTTACATCAACACCACCATTTACC 2637
Db	12 GCTACGAGGCGGCTCAAGTTCCTGGAGAGATTGGCTTACATCAACACCACCATTTACC 71
QY 2638	CACAACTCTCTCCCGCTTCTAGCTATGTTATATGCGCTATCTGTGCTCACTG 2697
Db	72 CACTAACCTCTCTCCCGCTTCTAGCTATGTTATATGCGCTATCTGTGCTCACTG 131
QY 2698	GAATGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTCGCTCTCC 2757
Db	132 GAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTCGCTCTCC 191
QY 2758	TTTCAATTTTCGCACATGGTATCTCTGAGATGAGGTGGAGTGGTGGCATTCACGAGT 2817
Db	192 TTTCAATTTTCGCACATGGTATCTCTGAGATGAGGTGGAGTGGTGGCATTCACGAGT 251
QY 2818	GGTGAGGAATGAACAGTCTGGGTATCTGGAGTATCTCTGCACATCTGTTGCCCTCT 2877
Db	252 GGTGAGGAATGAACAGTCTGGGTATCTGGAGTATCTCTGCCCATCTGTTGCCCTCT 311
QY 2878	TTACAGGTCTTCTGAAGGTGCTTCCCGGTATCTGCACACCAACTTCACCTGTCACCTCAAGG 2937
Db	312 TTCAGGTCTTCTGAAGGTGCTTCCCGGTATCTGCACACCAACTTCACCTGTCACCTCAAGG 371

QY 2938	CTAATGACGAAGAAGGCGACTT	750 bp mRNA linear EST 17-Oct-2001
Db	372 CTAATGATGAAGAAGGCGACTT	algare seedling shoot EST library
QY 2998	TCCCTCCGACGACCATTTTGA	Hordeum vulgare cDNA clone HVSMEA0017K05f,
Db	432 TCCCTCCGACGACCATTTTGA	mRNA sequence.
QY 3058	CCATCAACAGTGGTTACCAATC	BF624748
Db	492 CCATCAACAGTGGTTACCAATC	HVCDNA0001 (Cold str
QY 3118	GGGTGATTGTTCACTTATACCC	BF624748
Db	552 GGGTGAATGTTCACTTATACCC	BF624748.2 GI:13085
QY 3178	CGACGATTGTCATCGTCTGGGG	EST.
Db	612 CGACGATTGTCATCGTCTGGGG	Hordeum vulgare.
QY 3238	GTGTTGATCCATTACATACCCG	Hordeum vulgare
Db	672 GTGTTGATCCATTACATACCCG	EST.
RESULT 15		
LOCUS	BF624748	
DEFINITION	HVSMSEA0017K05f Horde	
ACCESSION	BF624748	
VERSION	BF624748.2	
KEYWORDS	EST.	
SOURCE	Hordeum vulgare.	
ORGANISM	Hordeum vulgare	
REFERENCE	1 (bases 1 to 750)	
AUTHORS	Wing, R., Close, T.J.,	
TITLE	Development of a gen	
JOURNAL	library	
COMMENT	Unpublished (2001)	
	On Dec 18, 2000 this	
	Contact: Wing RA	
	Clemson University	
	100 Jordan Hall, Cle	
	Tel: 864 656 7288	
	Fax: 864 656 4293	
	Email: rwing@clemson	
	Total hg bases = 554	
	Seq primer: AATTAACC	
	High quality sequenc	
FEATURES	Location/Qualifiers	
source	1..750	
	/organism="	
	/cultivar="	
	/db_xref="t	
	/clone="HVS	
	/clone_lib=	
	HVCDNA0001	
	/tissue_type	
	/lab_host="	
	/note="Vect	
	seeds were	
	conditions	
	with water,	



crystallization dishes. Five-day old seedlings were incubated at 50c for 2 days. Shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, and 60000 pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main ). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

BASE COUNT	212 a	148 c	195 g	
ORIGIN				
Query Match	19.0%; Score 687.8; DB 12; Length 750;			
Best Local Similarity	95.6%; Pred. No. 7.2e-181;			
Matches 718; Conservative	0; Mismatches 32; Indels 1; Gaps 1;			
QY	1479	GCATTGAAGTCCCGAGGAAGTGGATCATGCAAGATGGCACACCATGGCCAGGAAC	1538	
Db	1	GCATTGAAGTCCCGAGGAAGTGGATCATGCAAGATGGCACACCATGGCCAGGAAC	60	
QY	1539	AATACCAGGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAACAGTGGTGGCCTTGAT	1598	
Db	61	AATACCAGGATCATCCTGGAATGATTCAGGTTTTCCTTGGTCAACAGTGGTGGCCTTGAT	120	
QY	1599	ACTGAGGGTAATGAGTCCCGTTCCTTATGTCTCTCGTGAAGGGTCTCGGTTC	1658	
Db	121	ACCGAGGGTAATGAGTCCCGTTCCTTATGTCTCTCGTGAAGGGTCTCGGTTC	180	
QY	1659	CAGCACCAAGAAGCTGGTGCATGAATGCCCTTGTTCGTCTCAGCTGTCTCTACT	1718	
Db	181	CAGCACCAAGAAGCTGGTGCATGAATGCCCTTGTTCGTCTCAGCTGTCTCTACT	240	
QY	1719	AATGGACAATACATGTTGAATCTTGATGTGATCACTACATCAACACAGCAAGGCTGTC	1778	
Db	241	AATGGACAATACATGTTGAATCTTGATGTGATCACTACATCAACACAGCAAGGCTGTC	300	
QY	1779	CGAGAAGCTATGTCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAG	1838	
Db	301	CGGAAGCTATGTCTTCTTAATGGATCCAAACCTAGTCCGCAAGTCTGTTATGTGCAG	360	
QY	1839	TTCCCAACAAGGTTTTCATGGATTTGATAGGAATGATCGATATGCAACAGGAACACTGTC	1898	
Db	361	TTCCCAACAAGGTTTTCATGGATTTGATAGGAATGATCGATATGCAACAGGAACACTGTC	420	
QY	1899	TTTTTTGATATTAACTTGAAGGGCCTTGAGCGCATTCAGGACCCAGTTTATGTGGAACT	1958	
Db	421	TTTTTTGATATTAACTTGAAGGGCCTTGAGCGCATTCAGGACCCAGTTTATGTGGAACT	480	
QY	1959	GGTTGTGTTTTCACAGACAGCTATCTATGGTTATGAGCCCCCAATTAAAGCGAAGAAG	2018	
Db	481	GGTTGTGTTTTCACAGACAGCTATCTATGGTTATGAGCCCCCAATTAAAGCGAAGAAG	540	
QY	2019	CCAGGTTTCTTGGCATCACTATCTGCGGGCAAGAAGCAAGCAAGTCAAGAAAGG	2078	
Db	541	CCAAGTTTCTTGGCATCACTATCTGCGGGCAAGAAGCAAGCAAGTCAAGAAAGG	600	
QY	2079	AGCTCAGATAAGAAAAGTCGAACAGCATGTGGACAGTCTCTTCCAGTATTCATCTC	2138	
Db	601	AGCTCAGATAAGAAAAGTCGAACAGCATGTGGACAGTCTCTTCCAGTATTCATCTC	660	

QY	2139	GAAGACATAGACGAGGCTGTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATG	2198
Db	661	GAAGACATAGACGAGGCTGTTGAAGGTGCTGGGTTTGATGATGAGAAATCAGTTCTCATG	719
QY	2199	TCTCAAAATGAGCTTAGAGAGAGAGATTGGCC	2229
Db	720	TCTCAAAATGAACTTTAAAGAAAAAATTCGCC	750

Search completed: February 25, 2003, 19:51:29  
Job time : 4769 secs







Db	3213	CCACGAGATCGAGAACAAACACCGACGAGCTGAAGTTCAAGAACCCGCGAGGAGGAGGAGGT	3277
QY	113	GTGCCAGATCTGGCGGACGGGCTGGGCACCAACGCTTGACGCGGACACGTCTTTCACCGCGCTG	172
Db	3273	GTACC---CCACCGACCGGGCACCTGCAACGACTACACCGCCACCGAGGSCACCGCGCG	3329
QY	173	CGACGTCTGCGGTTCCTCGGTCTGCGGCGCTCTACGAGCACGACGCGCAAGAGGGAC	232
Db	3330	CTGCGCGGACGCTGCAACAGCGCAACGCGGCTTACGAGGACGCTTACGAGGTGGAC	3389
QY	233	CCAGGCGTGCCTTCCAGTGAACACCAACGTTACAAGCGCCACAGAGGGAGGCCCGCGGATCCG	292
Db	3390	CACCGCCACGCTGACTACAAGCCCACTACGAGGAGGACCTACACCGAGCTGCCCG	3449
QY	293	CGGGGAGGAAGCGCAGC	311
Db	3450	CGACACCACTCGGAGTAC	3468
RESULT 3			
US-08-459-448A-6			
; Sequence 6, Application US/08459448A			
; Patent No. 5859336			
; GENERAL INFORMATION:			
; APPLICANT: Koziel, Michael G.			
; APPLICANT: Desai, Nalini M.			
; APPLICANT: Lewis, Kelly S.			
; APPLICANT: Kramer, Vance C.			
; APPLICANT: Warren, Gregory W.			
; APPLICANT: Evola, Stephen V.			
; APPLICANT: Crossland, Lyle D.			
; APPLICANT: Wright, Martha S.			
; APPLICANT: Merlin, Ellis J.			
; APPLICANT: Launis, Karen L.			
; APPLICANT: Rothstein, Steven J.			
; APPLICANT: Bowman, Cindy G.			
; APPLICANT: Dawson, John L.			
; APPLICANT: Dunder, Erik M.			
; APPLICANT: Pace, Gary M.			
; APPLICANT: Suttie, Janet L.			
; TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED			
; NUMBER OF SEQUENCES: 94			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: No. 5859336artis Corporation			
; STREET: Patent & Trademark Dept., 520 White Plains			
; STREET: Rd., POB 2005			
; CITY: Tarrytown			
; STATE: New York			
; COUNTRY: USA			
; ZIP: 10591-9005			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/459,448A			
; FILING DATE: 02-JUN-1995			
; CLASSIFICATION: 800			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/951,715			
; FILING DATE: 25-SEP-1992			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US 07/772,027			
; FILING DATE: 04-OCT-1991			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Pace, Gary M.			
; REGISTRATION NUMBER: 40403			
; REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: (919)541-8582			
; TELEFAX: (919)541-8689			

Query Match	Best Local Similarity	1.68
Matches 141; Conservative	54.4%	
QY 53	CCGCGCATGGACGGCGACGGC	
Db 3213	CCAGGAGATCGAGACACACAC	
QY 113	GTGCCAGATCTGGCGGACGGG	
Db 3273	GTACC---CCACCGACACCGGG	
QY 173	CGACGTCTGCCGCTTCCCGGTT	
Db 3330	CTGCGCGGACGGCTTGCACAG	
QY 233	CCAGGCGCTGCTTCCAGTGCAG	
Db 3390	CACGGCAGCGTGACTACAA	
QY 293	CGGGGAGGAAGCGGACGAC	
Db 3450	CGACACCCTCGGAGTAC	
RESULT 5		
US-08-459-504B-6		
; Sequence 6: Application US/08		
; Patent No. 6075185		
; GENERAL INFORMATION:		
; APPLICANT: Kozziel, Michael		
; APPLICANT: Desai, Nalini		
; APPLICANT: Lewis, Kelly S		

[illegible]



APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttie, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 607518artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,504B  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/459,595  
FILING DATE: 02-JUN-1995  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC1577/CIP/DIV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3624 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Synthetic DNA"  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..3621  
OTHER INFORMATION: /product= "Full-length, maize  
OTHER INFORMATION: optimized cryiB"  
OTHER INFORMATION: /note= "Disclosed in Figure 6."

Query Match 1.6%; Score 59; DB 3; Length 3624;  
Best Local Similarity 54.4%; Pred. No. 6.6e-06;  
Matches 141; Conservative 0; Mismatches 115; Indels 3; Gaps 1;  
Qy 53 CCGCGCATGTGACGGCAGCGGACCCCTGAAGTCGGGAGGACGCGGCGCGGACGT 112  
Db 3213 CCACGAGATCGAACAACACCGAGAGCTGAAGTTCAAGACCGCGGAGGAGGAGT 3272  
Qy 113 GTGCCAGATCTGCGCCGACGCGCTGGGACACACGCTTGGACGGCGAGCTTTCACCGCGCT 172

Db 3273 GTACC---CCACCGACACCGGCACCTGCAACGACTACACCGCCACCGGCGCGG 3329  
Qy 173 CGAGCTGCGCGCTTCCGGTCTGCGGCCCTGTAGAGCAGCAGGAGGAGGAC 232  
Db 3330 CTGCGCGCAGCGCTGCAACAGCCGCGGCTACGAGGACGCTACGAGGTGGAC 3389  
Qy 233 CCAGGCTGCTCGAGTGAAGACCAAGTACAAGCGCACAGAGGGAGCCGCGATCCG 292  
Db 3390 CACGCGCAGCGTGAACCTACAAGCCACCTACGAGGAGGAGACCTACACCGAGTGC 3449  
Qy 293 CGGGGAGGAAGCGGACGAC 311  
Db 3450 CGACAACCACTGCGAGTAC 3468  
RESULT 6  
US-08-459-444-6  
; Sequence 6, Application US/08459444A  
; Patent No. 6121014  
; GENERAL INFORMATION:  
; APPLICANT: Koziel, Michael G.  
; Desai, Nalini M.  
; Lewis, Kelly S.  
; Kramer, Vance C.  
; Warren, Gregory W.  
; Evola, Stephen V.  
; Crossland, Lyle D.  
; Wright, Martha S.  
; Merlin, Ellis J.  
; Launis, Karen L.  
TITLE OF INVENTION: METHOD FOR PRODUCING A PLANT-OPTIMIZED  
NUCLEIC ACID CODING SEQUENCE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6121014artis Agribusiness Biotechnology Research, Inc.  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/459,444A  
FILING DATE: 02-Jun-1995  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/951,715  
FILING DATE: 25-SEP-1992  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: S-18805/P1/CGC1577/CIP/DIV6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8587  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3624 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: other nucleic acid  
DESCRIPTION: /desc = "Synthetic DNA"  
HYPOTHETICAL: NO  
FEATURE:  
NAME/KEY: CDS



```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: S-18805H
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8587
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Synthetic DNA"
; HYPOTHETICAL: NO
; FEATURE:
;
; LOCATION: 1..3621
; OTHER INFORMATION: /product= "Full-length, maize
; optimized cryiB"
; /note= "disclosed in Figure 6."
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-547-422-6

Query Match 1.6%; Score 59; DB 4; Length 3624;
Best Local Similarity 54.4%; Pred. No. 6.6e-06;
Matches 141; Conservative 0; Mismatches 115; Indels 3; Gaps 1;

QY 53 CCGCGCATGTGCGGCGGACGCGGACGCGCTGAACTCCGGGAGGCGGCGGGGACGT 112
Db 3213 CCACGAGATCGAACAACACCGACGAGCTGAAGTTCAGAACCGGAGGAGGAGGT 3272
QY 113 GTGCCAGATCTGCGCGGACGGCTGGGACACACGTTGGACGGAGCTTTCACCGCGTG 172
Db 3273 GTACC---CCACCGACACCGGACCTGCAACGACTACACCGCCACAGGGCGCGCGG 3329
QY 173 CGAGCTCTCGCGTTCGCGGCTGCGCGCCCTGCTACGAGCAGCAGCGCAAGGAGGCAC 232
Db 3330 CTGCGCGGACGCGCTGCAACAGCGGCAACGCGGCTACGAGGAGCGCTACGAGGTGGACAC 3389
QY 233 CCAGGCGCTGCTCCAGTGAAGACCAAGTACAAGCCGACAGAGGAGGAGGAGGATCG 292
Db 3390 CACCGCGGAGTGAATACAGGCCACCTACGAGGAGGAGACCTACACCGAGCTGGCGG 3449
QY 293 CGGGGAGGAGGCGGACGAC 311
Db 3450 CGACAACCACTGGGAGTAC 3468

RESULT 9
US-09-197-649-7
; Sequence 7, Application US/09197649
; Patent No. 6194550
; GENERAL INFORMATION:
; APPLICANT: Tuerk, Craig
; APPLICANT: Gold, Larry
; APPLICANT: Pribnow, David
; APPLICANT: Smith, Jonathan D.
; TITLE OF INVENTION: Systematic Polypeptide Evolution by Reverse Translation
; FILE REFERENCE: NEX02/C1-CON
; CURRENT APPLICATION NUMBER: US/09/197,649
; CURRENT FILING DATE: 1998-11-23
; EARLIER APPLICATION NUMBER: 07/829,461
; EARLIER FILING DATE: 1992-01-31
; EARLIER APPLICATION NUMBER: 07/739,055
; EARLIER FILING DATE: 1991-08-01
; EARLIER APPLICATION NUMBER: 07/561,968
; EARLIER FILING DATE: 1990-08-02
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 390
;
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Sequence
; OTHER INFORMATION: having a 120 repeat of ACG flanked by fixed
; OTHER INFORMATION: fragments having NcoI restriction sites.
US-09-197-649-7

Query Match 1.5%; Score 55.4; DB 4; Length 390;
Best Local Similarity 48.6%; Pred. No. 1.8e-05;
Matches 152; Conservative 0; Mismatches 161; Indels 0; Gaps 0;

QY 55 GCGCATGAGCGGCGGACGCGGACGCGCTGAACTCCGGGAGGCGGCGGGGACGTGT 114
Db 1 GGGCCATGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 60
QY 115 GCCAGATCTGCGCGGACGGCTGGGACCACTGGTGGACGGGAGCTTTCACCGCCTGGG 174
Db 61 ACGACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 120
QY 175 ACGTCTGCGGCTTCGCGGCTGCGCGCCCTGCTAGGACGACGAGCGCAAGGAGGCACCC 234
Db 121 ACGACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 180
QY 235 AGGCTGCTCCAGTGAAGACCAAGTACAAGCGCCACAGAGGAGGAGGAGGATCGCGG 294
Db 181 ACGAGGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 240
QY 295 GGGAGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 354
Db 241 ACGACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 300
QY 355 GCACGTGAGGACCA 367
Db 301 ACGACGACGAGCA 313

RESULT 10
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
```



```

1  APPLICANT:  Uknes, Scott Joseph
2  TITLE OF INVENTION:  Genes for the synthesis of
3  TITLE OF INVENTION:  antipathogenic substances
4  NUMBER OF SEQUENCES:  22
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE:  Ciba-Geigy Corporation
7  STREET:  7 Skyline Drive
8  CITY:  Hawthorne
9  STATE:  NY
10 COUNTRY:  USA
11 ZIP:  10532
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE:  Floppy disk
14 COMPUTER:  IBM PC compatible
15 OPERATING SYSTEM:  PC-DOS/MS-DOS
16 SOFTWARE:  Patentin Release #1.0, Version #1.25
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER:  US/08/457,342
19 FILING DATE:  01-JUN-1995
20 CLASSIFICATION:  424
21 PRIOR APPLICATION DATA:
22 APPLICATION NUMBER:  US 08/457,205
23 FILING DATE:  01-JUN-1995
24 APPLICATION NUMBER:  08/258,261
25 FILING DATE:  08-Jun-1994
26 ATTORNEY/AGENT INFORMATION:
27 NAME:  Elmer, James Scott
28 REGISTRATION NUMBER:  36,129
29 REFERENCE/DOCKET NUMBER:  CGC 1506/CIP3
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE:  919-541-8614
32 TELEFAX:  919-541-8689
33 INFORMATION FOR SEQ ID NO:  6:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH:  28958 base pairs
36 TYPE:  nucleic acid
37 STRANDEDNESS:  single
38 TOPOLOGY:  linear
39 MOLECULE TYPE:  DNA (genomic)
40 HYPOTHETICAL:  NO
41 ANTI-SENSE:  NO
42 US-08-457-342-6

```

```

; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-457-342-6

Query Match      1.5%; Score 53.4; DB 1; Length 28958;
Best Local Similarity 52.5%; Pred. No. 0.00072;
Matches 117; Conservative 0; Mismatches 106; Indels 0; Gaps 0;

QY  54  CGCGCCATGGACGGCGACGGCGACGCCCTGAAGTCCGGGAGGCACGGGGCCGGGGACGTG 113
      ||||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  24526  CGCGCCTTGGCGCCGACGGCGGCTCCAAGACCTTCTCGGACACGCGGCTACGGA 24585

QY  114  TGCACATCTGGCCGACGGGCGCTGGGACACAGTGTGACGGCGACGCTCTTCACCGCCTCG 173
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  24586  CGCGGAGAGGGCTGCTGCTGCTTCCCTCGAGCGGATCGGCGACGCCCTCGCCCGGAGA 24645

QY  174  GAGCTGTGCGGCTTCCCGTCTGCGCGCCCTGCTACGAGCACGAGCGCAAGGAGGCACC 233
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  24646  CACCGGCTCTGCTCTGTCGCGGCGACCGCCATCAACCAACGACGCGCGTCTGACGGT 24705

QY  234  CAGGCCTGCTCCAGTCAAGACCAAGTACAAGGCCACAGAG 276
      || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  24706  ATCACCGCCCCAACGGCACCTCCCGAGCAGAAGGTCTCTCCGGG 24748

RESULT 14
US-08-457-646A-6
; Sequence 6, Application US/08457646A
; Patent No. 5679560
; GENERAL INFORMATION:
; APPLICANT: Schnupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven

```

RESULT 14  
US-08-457-646A-6  
: Sequence 6, Application US/08457646A  
: Patent No. 5679560  
: GENERAL INFORMATION:  
: APPLICANT: Schupp, Thomas  
: APPLICANT: Ligon, James M.  
: APPLICANT: Beck, James Joseph  
: APPLICANT: Hill, Dwight Steven





Job time : 589 secs

\_\_\_\_\_



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2003, 18:32:29 ; Search time 181 Seconds  
(without alignments)  
11251.406 Million cell updates/sec

Title: US-09-900-237-29  
Perfect score: 3626  
Sequence: 1 gcacgaggaacccgctcca.....tcacattttgaggagtttt 3626

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 442118 seqs, 280819700 residues 884236

Total number of hits satisfying chosen parameters: 884236  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published Applications\_NA:\*
- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
  - 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
  - 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
  - 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
  - 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
  - 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
  - 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
  - 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
  - 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3626	100.0	3626	10	US-09-900-237-29
2	1681.6	46.4	1733	10	US-09-900-237-19
3	1364.4	37.6	3563	10	US-09-900-237-25
4	1340.2	37.0	3786	10	US-09-900-237-7
5	1320.4	36.4	3776	10	US-09-900-237-3
6	1271	35.1	3936	10	US-09-900-237-9
7	1257	34.7	3517	10	US-09-900-237-13
8	1215.6	33.5	3328	9	US-09-838-539-1
9	1192.6	32.9	3255	9	US-09-938-842A-559
10	1191.4	32.9	3255	9	US-09-938-842A-2530
11	1183.4	32.6	2890	10	US-09-900-237-17
12	1069.6	29.5	2663	10	US-09-900-237-23
13	966.8	26.7	2125	10	US-09-900-237-15
14	821.4	22.7	1560	10	US-09-900-237-27
15	524.8	14.5	1221	10	US-09-900-237-1
16	518.8	14.3	1189	10	US-09-900-237-5
17	419.6	11.6	1029	10	US-09-900-237-21
18	324.2	8.9	606	10	US-09-770-149-916
19	281.6	7.8	384	10	US-09-878-574-797

20	240	6.6	377	10	US-09-878-574-3023	Sequence 3023, Ap
21	202	5.6	283	10	US-09-923-876-4542	Sequence 4542, Ap
22	193.4	5.3	272	10	US-09-523-876-4807	Sequence 4807, Ap
23	187	5.2	271	10	US-09-878-574-11933	Sequence 11933, A
24	177.4	4.9	284	10	US-09-294-093B-1388	Sequence 1388, Ap
25	171.8	4.7	480	10	US-09-734-569-39	Sequence 39, Appl
26	171.8	4.7	1009	10	US-09-734-569-153	Sequence 153, App
27	167.2	4.6	255	10	US-09-878-574-5625	Sequence 5625, App
28	164.2	4.5	261	10	US-09-878-574-10197	Sequence 10197, A
29	160	4.4	278	10	US-09-294-093B-5812	Sequence 5812, Ap
30	157.8	4.4	280	10	US-09-878-574-14007	Sequence 14007, A
31	151.8	4.2	268	10	US-09-878-574-6337	Sequence 6337, Ap
32	149	4.1	259	10	US-09-878-574-13937	Sequence 13937, A
33	142.2	3.9	260	10	US-09-878-574-11087	Sequence 11087, A
34	138.6	3.8	496	10	US-09-734-569-37	Sequence 37, Appl
35	138	3.8	285	10	US-09-294-093B-4005	Sequence 4005, Ap
36	136.6	3.8	2268	10	US-09-887-576-669	Sequence 669, App
37	136.4	3.8	254	10	US-09-923-876-5625	Sequence 5625, Ap
38	134.6	3.7	258	10	US-09-923-876-2317	Sequence 2317, Ap
39	133.4	3.7	280	10	US-09-294-093B-1384	Sequence 1384, Ap
40	130	3.6	265	10	US-09-878-574-8805	Sequence 8805, Ap
41	128.4	3.5	474	10	US-09-770-444-100	Sequence 100, App
42	128.4	3.5	474	10	US-09-924-035A-217	Sequence 217, App
43	128	3.5	258	10	US-09-923-876-1909	Sequence 1909, Ap
44	127.4	3.5	226	10	US-09-294-093B-4855	Sequence 4855, Ap
45	125.8	3.5	2139	9	US-09-938-842A-2652	Sequence 2652, Ap

ALIGNMENTS

RESULT 1						
US-09-900-237-29						
; Sequence 29, Application US/09900237						
; Patent No. US20020120124A1						
; GENERAL INFORMATION:						
; APPLICANT: Allen, Stephen						
; TITLE OF INVENTION: Plant Cellulose Synthases						
; FILE REFERENCE: B81170 US CIP						
; CURRENT APPLICATION NUMBER: US/09/900,237						
; PRIOR FILING DATE: 2001-07-06						
; PRIOR APPLICATION NUMBER: 60/092,844						
; PRIOR FILING DATE: 1998-07-14						
; PRIOR APPLICATION NUMBER: PCT/US99/15871						
; PRIOR FILING DATE: 1999-07-13						
; PRIOR APPLICATION NUMBER: 09/720383						
; PRIOR FILING DATE: 2000-12-21						
; NUMBER OF SEQ ID NOS: 33						
; SOFTWARE: Microsoft Office 97						
; SEQ ID NO 29						
; LENGTH: 3626						
; TYPE: DNA						
; ORGANISM: Triticum aestivum						
US-09-900-237-29						
Query Match						
Best Local Similarity 100.0%; Score 3626; DB 10; Length 3626;						
Matches 3626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
QY	1	GCACGAGGAACCCGCTCCAGCTCTGTCGTCGGTCCGGGTTGGATFCGCTCTGCCGCGCCA	60			
Db	1	GCACGAGGAACCCGCTCCAGCTCTGTCGTCGGTCCGGGTTGGATFCGCTCTGCCGCGCCA	60			
QY	61	TGGAGCGGACGCGGACCCCTGAAGTCGGGAGGACGGGGCGGACCTGTGCCAGA	120			
Db	61	TGGAGCGGACGCGGACCCCTGAAGTCGGGAGGACGGGGCGGACCTGTGCCAGA	120			
QY	121	TCTGCGCGACGCGCTGGGACACACGCTTGGACGGGACGCTTACCCGCTGCCACGCT	180			
Db	121	TCTGCGCGACGCGCTGGGACACACGCTTGGACGGGACGCTTACCCGCTGCCACGCT	180			
QY	181	GCCGTTCCCGGTCGTGCCGCCCTCTACGAGCAGCGGCAAGGAGGCGCACCGGCT	240			
Db	181	GCCGTTCCCGGTCGTGCCGCCCTCTACGAGCAGCGGCAAGGAGGCGCACCGGCT	240			

Db 181 GCGCTTCCGGTGTGCGGCCCTGCTACGAGCAGCGCAAGGAGGCGCACCCAGGCCT 240  
QY 241 GCCTCCAGTCAAGCAACTACAAAGCGCCACAGAGGAGCCAGCGCATCCGGGGAGG 300  
Db 241 GCGCTCCAGTCAAGCAAGTACAAAGCGCCACAGAGGAGCCAGCGCATCCGGGGAGG 300  
QY 301 AAGCGCAGCACTGATGCGCGATGATGGTAGTGACTTCAACTACCCCTGCACTGTGGCACTG 360  
Db 301 AAGCGCAGCACTGATGCGCGATGATGGTAGTGACTTCAACTACCCCTGCACTGTGGCACTG 360  
QY 361 AGGACCAAGACAGAGATGCTGACAGGATGCGCAGCTGGCGCATGAACACCGGGGGCA 420  
Db 361 AGGACCAAGACAGAGATGCTGACAGGATGCGCAGCTGGCGCATGAACACCGGGGGCA 420  
QY 421 GTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACA 480  
Db 421 GTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCCTCTCCAAGTATGACA 480  
QY 481 GTGGAGAGATCCCTAGGGGATACGTCCTCTCAGTCAACCAAGCCAGATGTCAGGAGAAA 540  
Db 481 GTGGAGAGATCCCTAGGGGATACGTCCTCTCAGTCAACCAAGCCAGATGTCAGGAGAAA 540  
QY 541 TCCTGGAGCTTCGCCCTGATCATCACAATCCGTCAAGGGAGTTCCTCGGCAATG 600  
Db 541 TCCTGGAGCTTCGCCCTGATCATCACAATCCGTCAAGGGAGTTCCTCGGCAATG 600  
QY 601 CTCGGTTCCCTATGTGAATCATTACCAAAATCCGTCAAGGGAGTTCCTCGGCAATG 660  
Db 601 CTCGGTTCCCTATGTGAATCATTACCAAAATCCGTCAAGGGAGTTCCTCGGCAATG 660  
QY 661 GGAATGTTGGCTGGAAGAGAGATTGATGGCTGGAAAATGAAGCAGGAGGTCGGA 720  
Db 661 GGAATGTTGGCTGGAAGAGAGATTGATGGCTGGAAAATGAAGCAGGAGGTCGGA 720  
QY 721 TTCCCATGACTAATGGGACAGCATTCCTCCCTGAGGTCGGGCACTACTGACATCG 780  
Db 721 TTCCCATGACTAATGGGACAGCATTCCTCCCTGAGGTCGGGCACTACTGACATCG 780  
QY 781 ATGATCTACTGATCAACATGAAGACGCTTTACTGAAATGATGAACCTCGCCAGCCTC 840  
Db 781 ATGATCTACTGATCAACATGAAGACGCTTTACTGAAATGATGAACCTCGCCAGCCTC 840  
QY 841 TATCTAGAAAAGTCCCATGCTTCTCCCAAAATAAATCCCTACAGAAATGGTCAATGCTC 900  
Db 841 TATCTAGAAAAGTCCCATGCTTCTCCCAAAATAAATCCCTACAGAAATGGTCAATGCTC 900  
QY 901 TCGGTTGGTGTCTAAGCATCTTCTGCACTACCGCTCTCACAAATCCTGCGTAATG 960  
Db 901 TCGGTTGGTGTCTAAGCATCTTCTGCACTACCGCTCTCACAAATCCTGCGTAATG 960  
QY 961 CATACCCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTGCTTTATCTGGATAC 1020  
Db 961 CATACCCACTGTGGCTTTTATCTGTTATATGTGAGATTTGGTTGCTTTATCTGGATAC 1020  
QY 1021 TGGATCAGTTCGCGAAGTGTTCCTCAATCAACCGGGAGACCTACCTTTGATAGACTGGCTT 1080  
Db 1021 TGGATCAGTTCGCGAAGTGTTCCTCAATCAACCGGGAGACCTACCTTTGATAGACTGGCTT 1080  
QY 1081 TAAGGTATGACCGAGAGAGTGAACCGTCTCAGTTGGCTGTGTGACATATTTGTCAGTA 1140  
Db 1081 TAAGGTATGACCGAGAGAGTGAACCGTCTCAGTTGGCTGTGTGACATATTTGTCAGTA 1140  
QY 1141 CAGTCGACCCCTTGAAGGAGCCACCTATCGTCACTGCAACACTGTGCTATCCATTTCTTG 1200  
Db 1141 CAGTCGACCCCTTGAAGGAGCCACCTATCGTCACTGCAACACTGTGCTATCCATTTCTTG 1200  
QY 1201 CTGTTGATATCCCGTGGACAAGTCTCTTCTATGATCTGATGACGAGCTTCAATGTC 1260  
Db 1201 CTGTTGATATCCCGTGGACAAGTCTCTTCTATGATCTGATGACGAGCTTCAATGTC 1260  
QY 1261 TGACTTTTGACGATTTGGCTGAGACTTCAGAGTTTGTAGAGAAATGGGTACCATTTGTGA 1320  
Db 1261 TGACTTTTGACGATTTGGCTGAGACTTCAGAGTTTGTAGAGAAATGGGTACCATTTGTGA 1320

QY 1321 AGAAGTATGACATTGAACCCAG 1380  
Db 1321 AGAAGTATGACATTGAACCCAG 1380  
QY 1381 TGAAGACCAAGTCCAGCCTTC 1440  
Db 1381 TGAAGACCAAGTCCAGCCTTC 1440  
QY 1441 AAGNAATTAATAACAGGATAA 1500  
Db 1441 AAGNAATTAATAACAGGATAA 1500  
QY 1501 GATGGATCATGCAAGATGGCA 1560  
Db 1501 GATGGATCATGCAAGATGGCA 1560  
QY 1561 TGATTCAGGTTTCTCTTGGTCA 1620  
Db 1561 TGATTCAGGTTTCTCTTGGTCA 1620  
QY 1621 GTTTAGTTTATGTCTCTCGTG 1680  
Db 1621 GTTTAGTTTATGTCTCTCGTG 1680  
QY 1681 CCATGAATGCCCTTTGTTCTGTC 1740  
Db 1681 CCATGAATGCCCTTTGTTCTGTC 1740  
QY 1741 TTGATTTGATCACTACATCA 1800  
Db 1741 TTGATTTGATCACTACATCA 1800  
QY 1801 TGGATCAAACTAGGTCGCGCA 1860  
Db 1801 TGGATCAAACTAGGTCGCGCA 1860  
QY 1861 TTGATAGGAATGATCGATATG 1920  
Db 1861 TTGATAGGAATGATCGATATG 1920  
QY 1921 GCCTTGAGCGGATTCAGGAGC 1980  
Db 1921 GCCTTGAGCGGATTCAGGAGC 1980  
QY 1981 CTATCTATGTTATGAGCCCG 2040  
Db 1981 CTATCTATGTTATGAGCCCG 2040  
QY 2041 GTGGGGCAAGAGAGGCAAG 2100  
Db 2041 GTGGGGCAAGAGAGGCAAG 2100  
QY 2101 ACAAGATGTGGACAGTTCGT 2160  
Db 2101 ACAAGATGTGGACAGTTCGT 2160  
QY 2161 AAGTGTCTGGTTTGTGATG 2220  
Db 2161 AAGTGTCTGGTTTGTGATG 2220  
QY 2221 GATTTGGCCAGTACAGCAGAT 2280  
Db 2221 GATTTGGCCAGTACAGCAGAT 2280  
QY 2281 AGTCCTCCACTCCAGATCTCT 2340  
Db 2281 AGTCCTCCACTCCAGATCTCT 2340  
QY 2341 AGGACAAGTCTGAATGGGAG 2400  
Db 2341 AGGACAAGTCTGAATGGGAG 2400

QY 2401 TTCTTACTGGATTCAAGATGACGCAAGAGGCTGGCGTTTCAGTCTATTATTCGATGCCAAGC 2460  
Db 2401 TTCTTACTGGATTCAAGATGACGCAAGAGGCTGGCGTTTCAGTCTATTATTCGATGCCAAGC 2460  
QY 2461 GCCCAGCTTCAAGGATCTGCCCCCATCAATCTTTTCAGATGGTCTGAAACCAAGTCTGC 2520  
Db 2461 GCCCAGCTTCAAGGATCTGCCCCCATCAATCTTTTCAGATGGTCTGAAACCAAGTCTGC 2520  
QY 2521 GGTGGGCTCTCGGTTCTGTGTTGAAATTCCTTTAGCCGGGATGCGCCCTTATGGTATGGCT 2580  
Db 2521 GGTGGGCTCTCGGTTCTGTGTTGAAATTCCTTTAGCCGGGATGCGCCCTTATGGTATGGCT 2580  
QY 2581 ACGGAGGGCCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCAC 2640  
Db 2581 ACGGAGGGCCCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCACTTACCCAC 2640  
QY 2641 TAACCTCTCTCCGCTCTAGTCTATTTGATATTTGCTCTATCTCTGCTCACTGGAA 2700  
Db 2641 TAACCTCTCTCCGCTCTAGTCTATTTGATATTTGCTCTATCTCTGCTCACTGGAA 2700  
QY 2701 AGTTCAATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTCAATGCGCTCTCCCTTT 2760  
Db 2701 AGTTCAATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTCAATGCGCTCTCCCTTT 2760  
QY 2761 CAATTTTCGCCACTGATCTCTGATGAGTGGAGTGGTCTTGGCATTTGACGAGTGGT 2820  
Db 2761 CAATTTTCGCCACTGATCTCTGATGAGTGGAGTGGTCTTGGCATTTGACGAGTGGT 2820  
QY 2821 GGAGGAATGAACAGTTCTGGGTCATTGGAGGTATCTCTGCACATCTGTTGGCGTCTTTC 2880  
Db 2821 GGAGGAATGAACAGTTCTGGGTCATTGGAGGTATCTCTGCACATCTGTTGGCGTCTTTC 2880  
QY 2881 AGGGTCTCTGAGAGTCTGGCGGTATCGACACCACTTCACTGTCACCTCAAAAGGCA 2940  
Db 2881 AGGGTCTCTGAGAGTCTGGCGGTATCGACACCACTTCACTGTCACCTCAAAAGGCA 2940  
QY 2941 ATGACGAAGAAGCGACTTTGCTGAGCTCTACATGTTCAAGTGGAGAGCTTCTCATCC 3000  
Db 2941 ATGACGAAGAAGCGACTTTGCTGAGCTCTACATGTTCAAGTGGAGAGCTTCTCATCC 3000  
QY 3001 CTCGACGACCAATTTTGATCATTAAACATGTTGGTGTCTTGTGTCCTGTCACCTCTACGCCA 3060  
Db 3001 CTCGACGACCAATTTTGATCATTAAACATGTTGGTGTCTTGTGTCCTGTCACCTCTACGCCA 3060  
QY 3061 TCACACAGTGTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTCTTTGGCTCTCTGGG 3120  
Db 3061 TCACACAGTGTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTCTTTGGCTCTCTGGG 3120  
QY 3121 TGATTGTTCACTTATACCAATTCCTCAAGGCTCTTATGGCAGGCAAAACCCGACACCA 3180  
Db 3121 TGATTGTTCACTTATACCAATTCCTCAAGGCTCTTATGGCAGGCAAAACCCGACACCA 3180  
QY 3181 CGATTGTCATGCTGGGCTGCTCCTCGCTTCTATCTTCCCTTGTGCTGGTTCGCTG 3240  
Db 3181 CGATTGTCATGCTGGGCTGCTCCTCGCTTCTATCTTCCCTTGTGCTGGTTCGCTG 3240  
QY 3241 TTGATCCATTCACCTACCCGCTCGCTGGCCCAAAATATCAACACCTTGGCATCAACTGCT 3300  
Db 3241 TTGATCCATTCACCTACCCGCTCGCTGGCCCAAAATATCAACACCTTGGCATCAACTGCT 3300  
QY 3301 AGGAAAGTGGAGTTTGTAGACAGAGAAATATAACAGTGTATGAGCAACCAACCCGCGA 3360  
Db 3301 AGGAAAGTGGAGTTTGTAGACAGAGAAATATAACAGTGTATGAGCAACCAACCCGCGA 3360  
QY 3361 GCCAGAGATATTTATGTTGGGTTGTGAATTAATTAAGTGTGAGAAAGTTGTCAAAATTTG 3420  
Db 3361 GCCAGAGATATTTATGTTGGGTTGTGAATTAATTAAGTGTGAGAAAGTTGTCAAAATTTG 3420  
QY 3421 AGAAACACATTTGTAAATAGATGTAATAGACCATCTACCGTTTTCATGAGGTTAAGCTC 3480  
Db 3421 AGAAACACATTTGTAAATAGATGTAATAGACCATCTACCGTTTTCATGAGGTTAAGCTC 3480  
QY 3481 TTCTTTTGGGAACAAGGAATCTCATTGGTGAACCTTATAGGAATTTTTCCTATGAGGCA 3540

Db 3481 TTCCTTTTGGAAACAAGGAATCTCATTTGGTAAACCTATAGGAATTTTCTCATGAGGCA 3540  
QY 3541 CTTTGGATTAGGAATGGACCTATGAATCTTGTATTTATTTATTTATATATAAATTTATCC 3600  
Db 3541 CTTTGGATTAGGAATGGACCTATGAATCTTGTATTTATTTATTTATATAAATTTATCC 3600  
QY 3601 TGTCTTTCACATTTTGGAGGAGTTT 3626  
Db 3601 TGTCTTTCACATTTTGGAGGAGTTT 3626  
RESULT 2  
US-09-900-237-19  
; Sequence 19, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Syntheses  
; FILE REFERENCE: B01170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 19  
; LENGTH: 1733  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (262)  
US-09-900-237-19

Query Match 46.4%; Score 1681.6; DB 10; Length 1733;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 1696; Conservative 0; Mismatches 25; Indels 0; Gaps 0;  
QY 1782 GAAGCTATGTCTTCTCTTAATGATGATCCAAACCTAGTTCGGCAAGTCTGTATGTGCACTTC 1841  
Db 1 GAAGCTATGTCTTCTCTTAATGATGATCCAAACCTAGTTCGGCAAGTCTGTATGTGCACTTC 60  
QY 1842 CCACAAAGTTTGATGGATTGATAGGAATGATCGATATGCAACAGGAACACCTGTCTTT 1901  
Db 61 CCACAAAGTTTGATGGATTGATAGGAATGATCGATATGCAACAGGAACACCTGTCTTT 120  
QY 1902 TTTGATATTAACCTGAGGGGCTTTCAGCGCATTCGAAGCAAGTATGTGGGAACCTGGT 1961  
Db 121 TTTGATATTAACCTGAGGGGCTTTCAGCGCATTCGAAGCAAGTATGTGGGAACCTGGT 180  
QY 1962 TGTGTTTTCACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAGGCCAAGAGCCA 2021  
Db 181 TGTGTTTTCACAGAACAGCTATCTATGTTTATGAGCCCCCAATTAAGGCCAAGAGCCA 240  
QY 2022 GGTTCCTTGGCATCACTATGTGGGGCAAGGAAGGCAAGCAAGTCAAGAAAAGGAGC 2081  
Db 241 GGTTCCTTGGCATCACTATGTGGGGCAAGGAAGGCAAGCAAGTCAAGAAAAGGAGC 300  
QY 2082 TCAGATAAGAAAAGTTCGAACAGCATGTGGACAGTCTGTTCAGATTAATCAATCTCGAA 2141  
Db 301 TCAGATAAGAAAAGTTCGAACAGCATGTGGACAGTCTGTTCAGATTAATCAATCTCGAA 360  
QY 2142 GACATAGAGGAGGTGTTGAAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCTCATGCT 2201  
Db 361 GACATAGAGGAGGTGTTGAAGGTGCTGGGTTTGTATGATGAGAAATCAGTTCTCATGCT 420  
QY 2202 CAAATGAGCTTAGAGAAGAGATTTGGCCAGTCAGCAGCATTTGTTGCCCTCCACTCTGATG 2261

Db 421 CAATGAGCTTAGAGAAGATTTGGCCAGTCACGCCAATTTGTTGCTCCAGCTCTGATG 480  
QY 2262 GAATATGTTGGTGTCTCAGTCTCCACTCCAGAAATCTCTTTTGAAGAAGCTATCCAT 2321  
Db 481 GAATATGTTGGTGTCTCAGTCTCCACTCCAGAAATCTCTTTTGAAGAAGCTATCCAT 540  
QY 2322 GTCATAGTTGTGCTATGAGGACAAGTCTGAATGGGGAAGTCTGATGTTGGATCTAT 2381  
Db 541 GTCATAAGTTGTGCTATGAGGACAAGTCTGAATGGGGAAGTCTGATGTTGGATCTAT 600  
QY 2382 GGATCTGTACAGAAAGATATCTTACTGGATTCAAGATGCACGCAAGAGGCTGCGCTCA 2441  
Db 601 GGATCTGTACAGAAAGATATCTTAACTGGATTCAAGATGCACGCAAGAGGCTGCGCTCA 660  
QY 2442 GTCTATTGCATGCCAACGCCAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGAT 2501  
Db 661 ATCTATTGCATGCCAACGCCAGCTTTCAAGGGATCTGCCCCCATCAATCTTTTCAGAT 720  
QY 2502 CGTCTGAACCAAGTCTGGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGCCGGCAT 2561  
Db 721 CGTCTGAATCAAGTCTGGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCAGCCGGCAT 780  
QY 2562 TGCCCTCTATTGATGCTACGGAGGGCGCTCAAGTTCTGGAGAGATTCGCTTACATC 2621  
Db 781 TGCCCTCTATTGATGCTACGGAGGGCGCTCAAGTTCTGGAGAGATTCGCTTACATC 840  
QY 2622 AACACCACCTTTACCACCAACCTCTCTCCGCTTCTAGTCTATTTATATTTGCCCTGCT 2681  
Db 841 AACACCACCTTTACCACCAACCTCTCTCCGCTTCTAGTCTATTTATATTTGCCCTGCT 900  
QY 2682 ATCTGTCTGCTCACTGAAAGTTTCAATCGCCAGAGATTTAGCAACTTGGCCAGTATCTGG 2741  
Db 901 ATCTGTCTGCTCACTGAAAGTTTCAATCGCCAGAGATTTAGCAACTTGGCCAGTATCTGG 960  
QY 2742 TTCAATCGCTCTCTCTTCAATTTTCGCCCATCGTATCTTGGAGATGAGTGAGTGGT 2801  
Db 961 TTCAATCGCTCTCTCTTCAATTTTCGCCCATCGTATCTTGGAGATGAGTGAGTGGT 1020  
QY 2802 GTTGGCAATTCAGGAGTGGTGGAGGAATGAACAGTCTCTGGTCAATTTGGAGGATCTCTGCA 2861  
Db 1021 GTTGGCAATTCAGGAGTGGTGGAGGAATGAACAGTCTCTGGTCAATTTGGAGGATCTCTGCC 1080  
QY 2862 CATCTGTTTGGCTTTTTCAGGCTCTCTGAGGCTTTCGCGGTATCGACACCAACTTC 2921  
Db 1081 CATCTGTTTGGCTTTTTCAGGCTCTCTGAGGCTTTCGCGGTATCGACACCAACTTC 1140  
QY 2922 ACTGTCACTCAAGGCTAATGATGAAGAGGCGACTTTGCTGAGCTCTACATGTTCAAG 2981  
Db 1141 ACTGTCACTCAAGGCTAATGATGAAGAGGCGACTTTGCTGAGCTCTACATGTTCAAG 1200  
QY 2982 TGGACAGCGTTCTATCCCTCCGACGACCAATTTTGATCAATTAACATGGTGTGTCGTT 3041  
Db 1201 TGGACAGCGTTCTATCCCTCCGACGACCAATTTTGATCAATTAACATGGTGTGTCGTT 1260  
QY 3042 GCTGGCACTCTACGCCATCAACAGTGTACCAATCATGGGGCGCTTTTGGGAAG 3101  
Db 1261 GCTGGTACCTCTACGCCATCAACAGTGTACCAATCATGGGGCGCTTTTGGGAAG 1320  
QY 3102 CTCTTCTTTTGGCTTCTGGGTGATTTGCTTATACCATTCCTCAAGGGTCTTATGGGC 3161  
Db 1321 CTCTTCTTTTGGCTTCTGGGTGATTTGCTTATACCATTCCTCAAGGGTCTTATGGGC 1380  
QY 3162 AGGCAAAACCGCACACGAGATTGTCATCGTCTGGGCTGTCTCTCGCTTCTATCTTC 3221  
Db 1381 AGGCAAAACCGCACACGAGATTGTCATCGTCTGGGCTGTCTCTCGCTTCTATCTTC 1440  
QY 3222 TCCTTGTGTGGGTCTGCTGTGATCCATTCACACCGCTCTCGTGGCCCAAAATATCCAA 3281  
Db 1441 TCCTTGTGTGGGTCTGCTGTGATCCATTCACACCGCTCTCGTGGCCCAAAATATCCAA 1500  
QY 3282 ACCTGTGGCATCACTGCTAGGAAAGTGGAGTTGTAGAGACAGAAAAATATAACAGTGA 3341  
Db 1501 ACCTGTGGCATCACTGCTAGGAAAGTGGAGTTGTAGAGACAGAAAAATATAACAGTGA 1560

QY 3342 TCGAGCAACACCCGCGGAGCG 3401  
Db 1561 TCGAGCGCACCACTGTGGAGCG 1620  
QY 3402 GAGNAAGTTGTCAAAATTGAG 3461  
Db 1621 GAGNAAGTTGTCAAAATTGAG 1680  
QY 3462 TTTTCATGAGTTAAAGCTCTTC 1721  
Db 1681 TTTTCATGAGTTAAAGCTCTTC 1780

RESULT 3  
US-09-900-237-25  
; Sequence 25, Application US/09-900-237-25  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cell Culture Media  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: 09/00,237  
; PRIOR FILING DATE: 2001-07-14  
; PRIOR APPLICATION NUMBER: 60/134,344  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US01/015871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/00,237  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 25  
; LENGTH: 3563  
; TYPE: DNA  
; ORGANISM: Impatiens balsamina  
US-09-900-237-25

Query Match 37.6%  
Best Local Similarity 66.2%  
Matches 2108; Conservative

QY 81 CTGAAGTCCGGGAGGCGACGGG 140  
Db 88 CCGAAGCCCTTAAAGAGGCTT 147  
QY 141 ACCACGTTGGACGCGGACGCT 200  
Db 148 AATCAGCCACCGCGGACACT 207  
QY 201 CCCTGTACGAGCACGAGCGCA 260  
Db 208 CCTTGTACGAGTACGAAAGGA 267  
QY 261 TACAAGCGCCACAGAGGGAGC 320  
Db 268 TACAAAAGACAGAAAGGAGTC 327  
QY 321 GATGATGGTGTAGTCTTCAACT 380  
Db 328 GATTTGAAAACGAGTTCAAT 384  
QY 381 GCTGACAGGATGGCGAGCTGGC 440  
Db 385 ACGCAAGCGGGCCATGGCAGG 444  
QY 441 AAGTATGACAGTGGCGAGATCG 500  
Db 445 AGCATGATGAATCCCAACAA 504  
QY 501 TAGTCCCTTTCAGTCAACCAAC 560  
Db 505 GAAATTC--CACTCCCGATAA 562

Score 1364.4; DB 10; Length 3563;  
Id. No. 0;  
Mismatches 1026; Indels 48; Gaps 8;

CGAGTGTGCCAGATCTGCGCGGACGGCTGGC 140  
ACAGATATGTCAGATATGCGGTGACACAGTCGA 147  
CGCTGCGACGCTCTGCGGCTTCCCGGCTGCGCG 200  
CGCTGTAATAAATGTGGATTCCCGGTTGCGCG 207  
GGCACCCAGCGCTGCCCTCCAGTCAAGACCAAG 260  
GGAAACCAATGCTGCCCTCAGTCAAGACCAAGA 267  
ATCCGCGGGGAGGAGGAGGAGGAGTGTGGAT 320  
GTTGAAGGAGATGAAGAAGAGGAGGATGTGGAT 327  
GCATCTGGCACTGAGGACCAAGAGAGATCCCTAGGGA 380  
GGTAAGGGA--AGAACAGAGAGAGTAACC 384  
AACACCGGGGGAGTGGCAATGTTGGCCACCCC 440  
CAGCAGGATATTGAGCTCTCTGTTTCATCATCT 444  
TCCCAAGTATGACAGTGGAGAGATCCCTAGGGA 500  
CCCTTCTCTACTCAGGGGCACTCGGTATCGGGC 504  
ATGTCAGGAGAAATCCCTGGAGCTTCGCTGAT 560  
CTATAAGGACTACATCAGGTGCCCATAGGCCCTG 562



QY 561 CATCATGATGATCCCTACGGGGAACATCAGCAGACGTGCTCCGTTCCCTATGTGAAT 620  
Db 563 TGGAAAAATCCATTCCTCTATATCGATCCAGGACGACGTTCTGTGAGAAATATTTG 622  
QY 621 CATTCACCAATCCGTCACAGGAGTTCTCCGGCAGTATTGGAAATGTTCCCTGGAAAGAG 680  
Db 623 ACCCTTCARAGGACTTGAA-----CTCGTACGGCTTGGCAATGTGATTTGGAAGAA 675  
QY 681 AGATTGTAGTGGTGGAAAAATGAAGCAGGACAAAGGTGCGATTCCCATGACTAATGGACA 740  
Db 676 AGGGTTGAAGGTTGGAAACTTAAGCAGGAGAAAAATATGGTGCAATATGAC----- 725  
QY 741 AGCATTCCTCCCTCTCAAGTCGGCAGCTACTGCATGCATGATGCATCTACTGAATACAC 800  
Db 726 --CAGTAGATATCCTGAAGGGAAAGGAGATCTGAAGGTACGGGATCAATGGGAGGAA 783  
QY 801 ATGGAAGAGCGTTTACTGAATGATGAACCTCGCCAGCCCTCTATCTAGAAAAAGTCCCCATT 860  
Db 784 CTTCAAAATGGCTGCC-----GACGATATCCGACACACCATGAGCCGAATCGTCCCAT 837  
QY 861 GCTTCCTCCAAAATAAATCCCTACAGAAATGGTCAATGTTCTCGGGTGGTGTCTCTAAGC 920  
Db 838 TCGTCGACCGACCTCACTCCCTACAGAGTGGTTATCATACTTCGGTTAATATCTCTCGGT 897  
QY 921 ATCTTCCTGCACTACCGTCTCAAAATCCTGTGCGTAAATGCATACCCACTGTGGCTTTA 980  
Db 898 TTCTTCCTGCAATACCGTTGTACTCATCCAGTGAAGATGCTTATCCATTATGGCTTACC 957  
QY 981 TCTGTTATATGTGAGATTTGGTTTGTCTTATCCTGGATACTGGATCAGTTCCCGAAGTGG 1040  
Db 958 TCGGTTATTTGTGAAGTTTGGTTTGCATTGTCATGSCCTGCTTGATCAGTTCCCTTAATGG 1017  
QY 1041 TTTCCAATCAACGGGAGACCTACCTTGATAGACATGGCTTTAAGGTATGACCGAGAAGT 1100  
Db 1018 TCGCCCTGTGAACCGCGAGACTTATCTCGACAGACTGTCCATGAGATTCGATAGGAAGGG 1077  
QY 1101 GAACCGTCTCAGTGTGCTGTGTGACATATTTGTGACGTACAGTCGACCCCTTTGAAGGAG 1160  
Db 1078 GAGCCCTTCGAATTGGCGCAATTGACGTATTTGTGTAGTACCGTGGATCCTTTGAAGAG 1137  
QY 1161 CCACCTATCGTCTACTGCCAACACTGTGCTATCCATCTTGTGTTGATTATCCCGTGGAC 1220  
Db 1138 CCACCACTGTGACAGCCAAACACGGTTTGTCTATCCTGGCGTGGATTACCCGTGTTGAC 1197  
QY 1221 AAGGTCTCTGTCTATCTATGATGACGGAGCTTCAATGCTGACITTTGACCGCTTGGCT 1280  
Db 1198 AAGTCTCTGTCTATCTTTCGGATGACGGTTGAGCAATGTTGACCTTCGAAGCTCTATCC 1257  
QY 1281 GAGACTTCAGAGTTTCTAGGAAATGGGTACCATTGTGGAAGAAGTATGACATTGAACCC 1340  
Db 1258 GAGACAGCCGAGTTTCTGAAGAAATGGGCACCCTTCTGTGAAGAACATAGTATTGACCT 1317  
QY 1341 AGAGTCCCGAGTTTACTTTTGGCAGAAATTTGATTACCTGAAGACAAAGTCCAGCCT 1400  
Db 1318 CGGGCCCGCGAAATTTTATTTCCGCTCAAAAATTTGATTACTTGAAGGATAAGGTGACGCT 1377  
QY 1401 TCATTTTGTAAAGACCGCGGCGCATGAAGAGAGAAATATGAAGAATTTAAATTCAGGATA 1460  
Db 1378 TCTTTTCGTGAAGAGCGGAGGCGGATGAAGAGGGAATACGAAGAATTTAAGGTTAGGATT 1437  
QY 1461 AATGCCCTAGTTTCTAAGGCATTGAAAGTCCCGAGGAAGGATGGAATCATGCAAGATGCC 1520  
Db 1438 AATGCCGTTGTGCGAAGCGGCAAAAGTGCAGAAAGGATGGACGATGAAGATGGA 1497  
QY 1521 ACACATGGCCAGGAACAAATACAGGATCATCTCGGAATGATTCAGGTTTTCCTTGGT 1580  
Db 1498 ACTCCATGGCCGGAATAACTCGAGAGATCATCTCGGAATGATTCAGGTTTTTTTAGCC 1557  
QY 1581 CACAGTGGTGGCTTGATACTAGGGTAATAGCTCCCGGTTTAGTTTATGTGCTCGT 1640  
Db 1558 CATAGTGGGGGTTTCGATACGAGGGAATGAGTTACCTCCGCTGGTGTACGTTTCTCGT 1617

QY 1641 GAAAGCGTCTCTGGTTCACGACACACAAGAGGCTGGTGCCATCAATGCCCTTGTTCGT 1700  
Db 1618 GAGAAACGTCCTGGATTTCAGCATCACAAAGAACCGGGGCAATGAAGCGCATTTGATTCGA 1677  
QY 1701 GTCTCAGCTGCTCTACTTAATTTGGCAATACATCTTTGAATCTTGATTTGTGATCACTACATC 1760  
Db 1678 GTATCGGAGTGTCTGACAAATGGGGCTTATCTGCTTAACGTGGATTGTGATCACTACTTTC 1737  
QY 1761 AACACAGCAAGGCTGTCCGAGAAGCTATGTCTTCCTAATGATGATCAAAACCTAGGTCCG 1820  
Db 1738 AACACAGCAAAATGCTAAAAGAGGCAATGTGCTTTATGATGATGATCCTCAAACTTTGGAAG 1797  
QY 1821 CAAGTCTCTTATGTGCAAGTTCACACAGAGTTTGTATGGGATGATAGAAATGATCGATAT 1880  
Db 1798 AAAACATGTTACGTTACGTTCCCTCAACGGTTTGTATGATTTGACATTGCGAGATCGATAT 1857  
QY 1881 GCAAACAGAACACTGTCTTTTTTATATTAACTTTAGGGGCTTTCGGCGATTTCACAGGA 1940  
Db 1858 GCTAACCGTAACATTTCTTCTCGATATCAACTTTGAAAGGTTTGGACGGCATTCAGGGC 1917  
QY 1941 CAGTTTATGTGGAACTGGTTGTGTTTCAACAGACAGCTATCTATGTTTATGAGGCC 2000  
Db 1918 CCAGTTTATGTGGTACCGGTTGTGTTTCAACAGACAGCGCTATACGGGTATGATCCA 1977  
QY 2001 CCAATTAAAGGCGAAGAA-----GCCAGGTTTCTTGGCATCATTATGTGGGGCAAG 2051  
Db 1978 GTCTTAACAGAGGAGGATTTGGAACCGAATATCATCATCAAGAGCTGTTGCGGCTCGAGG 2037  
QY 2052 AAGAAGGCAACAAGTCAAGAAAAAGGAGCTCAGATAGAAAAAGTGCAGAACAGCATGTG 2111  
Db 2038 AAAAAGGTAAGGTGGCAACAAGAAGTACATTGCAAAAAACAGACACTAAACGGAACC 2097  
QY 2112 GACAGTCTCTGTTCAGATTTCATCTCGAAGACATAGAGAGGGTGTGTAAGGTGCTGGG 2171  
Db 2098 GAATCAACCGCGCCCATTTTCAATATGGAAGATATTGAAGAGGCAATTGAAGGT----- 2151  
QY 2172 TTTGATGATGAGAAATCAGTTCTCATGTCTCAAAATGAGCTTAGAAGAGAGATTGGCCAG 2231  
Db 2152 TACGATGACGAGAGATCTTTTCTCATGGCACA--GAGTTACGAAGAAGCGGTTGCGTCAA 2208  
QY 2232 TCAGCAGCATTTGTTCCTCCACTCTGATGGAATATGTTGTTCTTCAGTCTCTCCACT 2291  
Db 2209 TCCCTGTCTTATTGTGTCACGCTTCATGGAACAAGSGCGCTTCTCCTTCCACAAAC 2268  
QY 2292 CCAGNATCTCTTTGAAAGAGACTTATCCATGTCTATAAGTTGTGCTATGAGGACAAAGTCT 2351  
Db 2269 TCTGCAACCCCTTTGAAAGAGCAATCCATGTTATTAGCTGTGGGTACGAGGACAAGACT 2328  
QY 2352 GAATGGGGAACCTGAGATTGGTTGGATCTATGGATCTGTACAGAGAATATTCTTACTGGA 2411  
Db 2329 GAATGGGCAAGAGATTGGATGGATATATGGATCTGTAAAGGAAGATATCTTGACCGGG 2388  
QY 2412 TTCAGATGTCACGCAAGAGGCTGGCGTTTCACTTATTGTCATGCCCAAGCGCCAGCTTTC 2471  
Db 2389 TTCAGATGTCATACGAGAGGATGGATTCAATCTACTATTAATGTTTATCCACTCACCTCCAT 2448  
QY 2472 AAGGATCTGCCCCATCAATCTTTTCAGATCGTCTGAACCAAGTGTGCGGTGGGCTCTC 2531  
Db 2449 AAGGATCTGCCCATTAATCTTTTCGGAGCTGTTTGAACACAGGCTTTCGATGGGCTCTC 2508  
QY 2532 GGTTCGTGTGAATTTCTTTTCAGCGGCTTTCGCCCTTATGCTATGGCTACGAGGAGGCGC 2591  
Db 2509 GGATCGATTGAGATCTTTTGGATGAGATTTGCCCATTTGGTATGGCTACAGCGGTAGA 2568  
QY 2592 CTCAGTTTCTGGAGAGATTGCTTACATCAACACCACTTTTACCCACTAACCTCTCTC 2651  
Db 2569 CTCAGTTTCTGGAGAGATTGCTTATATCAATACTATTGTTTATCCACTCACCTCCAT 2628  
QY 2652 CGGCTTCTAGTCTATTTATGTTATATGCTGCTATCTGCTGCTCACTGGGAAAGTTTATCATG 2711  
Db 2629 CTTTACTTGTCTATTGCAACCTTCCCTGCTATCTGCTTACTTACCGGAAAGTTTATCTGTT 2688  
QY 2712 CCAGAGATTAGCAACTTGGCCAGTATCTGGGTTCTATGCGCTCTTCTTCAATTTTCGCC 2771

Db	2699	CCGGAGATAAGCAACTACGCGAGCATCGGTTCATCTCTCTCGTGCTATTTTCTCG	2748
Qy	2772	ACTGATATCCCTTGACATGAGGTGGAGTGGTGTGGCATTTGACGAGTGGTGGAGGAATGAA	2831
Db	2749	ACGGGAATACGGACCTAAGATGGAGCGGGTTACACTGGAGGACTGGTGGAGAAACGAG	2808
Qy	2832	CAGTCTCGGGTCATTGGAGGTATCTTCACATCTGTTTGGCGTCTTTCAGGGTCTTCG	2891
Db	2809	CAATTCTGGGTAATCGGTGGCAGTCGGCTCATCTCTTTGCGGTGTCCAAGSCCTGCTA	2868
Qy	2892	AAGTGTCTGGGGTATCGACACCACTTCACTGTCACTCAAAAGCTAATGACGAAGA	2951
Db	2869	AAAGTGTCTGGGGGATCGACACGAATTTACCGTCAGTTCGAAAGCGTCGGACGAGGAC	2928
Qy	2952	GGCAGCTTGTGAGCTCTACATGTTCAAGTGGAGAGCTTCTCATCCCTCCGAGCACC	3011
Db	2929	GGGACTTTGGGGAGCTTTAGCTTTTCAAGTGGAGTTCCTCTCTCATCCCTCCGACCA	2988
Qy	3012	ATTTTGATCATTAACATGGTGGTGTGCTGTGGCAGCTCTCAGCCATCAACAGTGTG	3071
Db	2989	ATCTGGTGTGAACATGGTGGGATAGTGCCGCGCTCTCGTTCGCATCAACAGTGA	3048
Qy	3072	TACCAATCATGGGGCCGCTCTTTGGGAAGCTTCTTTTGGCTTTGGGTGATGTTTCAC	3131
Db	3049	TACCAGTCTGGGAGCCGCTCTTCGGAAGTGTGTTCTTTGGCATATGGGTATTGTCCAT	3108
Qy	3132	TTATACCATTCCTCAAGGCTTTATGGCAGGCAAAACCCGACACCCGACGATTGTTCATC	3191
Db	3109	TTGTACCATTCCTTAAAGGTTGTGTGGAGCGGAGAAATCGGACACCCACGATTGTTAT	3168
Qy	3192	GTCTGGGCTGCTCTCTCGCTCTATCTTCTCTGCTGTGGTTCGTTCATCCATTC	3251
Db	3169	GTCTGGTCTGTACTTTTGGCTTCCATAATTTCTCTCTATGGGTGGTATGATCGGTTT	3228
Qy	3252	AC 3253	
Db	3229	AC 3230	
RESULT 4			
US-09-900-237-7			
; Sequence 7, Application US/09900237			
; Patent No. US20020120124A1			
; GENERAL INFORMATION:			
; APPLICANT: Allen, Stephen			
; TITLE OF INVENTION: Plant Cellulose Synthases			
; FILE REFERENCE: BB1170 US CIP			
; CURRENT APPLICATION NUMBER: US/09/900,237			
; CURRENT FILING DATE: 2001-07-06			
; PRIOR APPLICATION NUMBER: 60/092,844			
; PRIOR FILING DATE: 1998-07-14			
; PRIOR APPLICATION NUMBER: PCT/US99/15871			
; PRIOR FILING DATE: 1999-07-13			
; PRIOR APPLICATION NUMBER: 09/720383			
; PRIOR FILING DATE: 2000-12-21			
; NUMBER OF SEQ ID NOS: 33			
; SOFTWARE: Microsoft Office 97			
; SEQ ID NO 7			
; LENGTH: 3786			
; TYPE: DNA			
; ORGANISM: Zea mays			
; US-09-900-237-7			

	Query Match	37.0%	Score 1340.2	DB 10	Length 3786
	Best Local Similarity	65.5%	Pred. No. 0		
	Matches 2137	Conservative 0	Mismatches 1023	Indels 103	Gaps 8
QY	89	CGGAGGACACGGGCGGGGACGTTGTGCCATCTGCGCCAGCGCCCTGGGCACCACTTT	148		
Db	310	CGCGCGCGCGCGGCGGAGGCGCGTGCAGATATCGCGGACAGAGTCGGGGTGGGTT	369		
QY	149	GGAGGGGACGCTCTTACCGCGCTCGCGAGCTCGCGGCTCTCCGGGTCCTCGCGCCCTGCTA	208		

[illegible]

Db 1394 GTCTCCTGCTATGTTTCTGATGGTGTGCAATGCTTACGTTTGAAGCATTTGCTGAA 1453  
Qy 1284 ACTTCAGAGTTTGGTAGGAATGGGTACCAATTTGTGAAGAAGTATGACATTTGAACCCAGA 1343  
Db 1454 ACATCTGAATTTGCAAGAAATGGGTCTCTTCAGCAAAAGTTTAAATATCGAGCCCTGT 1513  
Qy 1344 GCTCCCGAGTTTACTTTTCCAGAAAATTTGATTACTTGAAGCAAAAGTCCAGCCCTTCA 1403  
Db 1514 GCTCCTGAGTGTACTTCCACAGAGATAGACTACTGAAAGACAAGGTGCTGCTTCA 1573  
Qy 1404 TTTGTTAAAGACCCCGGCCATGAAGAGAGAATATGAAGAATTTAAATACAGGATAAAT 1463  
Db 1574 TTTTGTAGGAGAGAGGCGATGAAGAGAAATACAGGAATTTCAAGGTAAAGATCAAT 1633  
Qy 1464 GCCCTAGTTTCTAAGGCATTTGAAGTCCCGAGGAGGATGATCATGCAAGATGCACA 1523  
Db 1634 GCCTTGTTGCAAAAGCCCAAGAGTTCTTGAGGAGAGGTGGCAATGCAAGATGGAAGC 1693  
Qy 1524 CCATGGCCAGGAACAAATACAGGGATCATCTTGGAAATGATTCAGGTTTTTCTTTGGTCCAC 1583  
Db 1694 CCCTGGCCTTGAACAACAGTACGCGATCATCTTGGAAATGATTCAGGTATTTCTTTGGCCAA 1753  
Qy 1584 AGTGGTGGCCTTGATCTAGAGGTAAATGAGCTCCCGGTTAGTTTATGTGTCCTGTGAA 1643  
Db 1754 AGTGGCGGTGCTGATGTGGAAGAAATGAGTTGCCCTGGCTTTATGTCTCGAGAGAA 1813  
Qy 1644 AAGCGTCTGGGTTCCAGACACCACAAGAGGCTGGTGCCATGAATGCCCTTTGTCGTGTC 1703  
Db 1814 AAGAGCCAGGTTATAACCATCACAGAGGCTGGTGCCATGAATGCATGCTGCTGTC 1873  
Qy 1704 TCAGCTGCTCTTACTAATGACATACATGTTGAAATCTTGAATGTGATCACTACATCAAC 1763  
Db 1874 TCTGCTGCTTATCAAAATGCTGATACCTATTGAACTTGGACTGTGATCACTACATCAAC 1933  
Qy 1764 AACAGAGGCTGTCGAGAGCTATGCTCTCTTATGGATCCAAACCTAGTCCGCGAA 1823  
Db 1934 AATAGCAAGGCCATAAAGAGGCTATGTGTTTCATATGATGATCCTTTGGTGGGGAAGAA 1993  
Qy 1824 GTCTGTTATGTGAGTTTCCACAAAAGTTTGAATGGATTTGATAGGAATGATCATATGCA 1883  
Db 1994 GTGTGCTATGTACAGTTCCCTCAGAGGTTTGTGATTTGATGACAAAATGATCATAGCT 2053  
Qy 1884 AACAGGAACACTGCTTTTTTGTATTAATCTTGAAGGGCTTGAAGGCAATCAAGACCA 1943  
Db 2054 AACAGSAACGTTGTCTTTTTTGACATCAACATGAAAGGTTTGGACGGTATTTCAAGGACCC 2113  
Qy 1944 GTTTATCTGGAACTGCTGTTTCAACAGAACAGCTATCTATGTTATGAGCCCA 2003  
Db 2114 ATTTATGTGGTACTGATGTGTTTTCAGACGCGCACTGTATGTTATGATGCTCCT 2173  
Qy 2004 ATTAAGCGCAAGACCA-----GGTTTCTTTGGCATCA 2036  
Db 2174 AAAACGAAGAAGCCACCATCAAGAACTTGCACTGCTGGCCCAAGTGGTGCCTCTTTCG 2233  
Qy 2037 CTATGTGGGGCAAGAGAGGCAAGCAAGTCAAGAAAGAGGCTCAGATAGAAAAA- 2095  
Db 2234 TGCTGACAGGAAACAGAAATAAAGAGAACTACAAAACCAAGACGGAAGAGAA 2293  
Qy 2096 -----GTCGAACAGCATGTGGACAGTTCTGTTCAGTATTTCAATCTCGAAGACATAGAG 2150  
Db 2294 AGATTATTTTCAAGAAAGCAGAAACCCATCTCTGCAATATGCTTTGGGTGAATTTGAT 2353  
Qy 2151 GAGGGTTTGAAGGTGCTGGGTTTGTATGATGAAATCAGTTCTCATGTCTCAAAATGAGC 2210  
Db 2354 GAAGGTGCTCCAGGTGC-----TGATATCGAAGAGCCGGAATCGTAAATCAACAGAAA 2407  
Qy 2211 TTAGAGAGAGATTTGGCCAGTCAGAGCATTTGTTGCCCTCCACTCTGATGGAATATGTT 2270  
Db 2408 CTAGAGAGAGAAATTTGGGCAGTCTTCTGTTTTTGTGCGCATCAACACTTCTTGAGAACGGA 2467  
Qy 2271 GGTGTTCTCAGTCTCCACTCCAGAAATCTCTTTTCAAGAAAGCTATCCATGCTCATAGT 2330  
Db 2468 GGGACCTGAAAGCCGCAAGTCCAGGTTCTCTTCTGAAGAAAGCTATACATGTTATCAGC 2527

Qy 2331 TGTGCTATGAGGACAAAGTCTGAATGGGAACTGAGATTTGGATCTATGGATCTGTGTC 2390  
Db 2528 TCGCGCTACGAAGACAAAGACCGACTGGGAAAAGAGATTTGGCTGGATTTACGGATCGATC 2587  
Qy 2391 ACAGAAGATATTCTTACTGGATTCAGAGTGCACGCAAGAGGCTGGCGTTCACTATTGTC 2450  
Db 2588 ACAGAGGATATCTTGACTTGGATTTAAGATGCATGCCATGGCTGGCGGTCTATTACTGC 2647  
Qy 2451 ATGCCAAAGCCGACAGCTTCAAGGATCTCCCCCAATCAATCTTTCAAGTCGCTGAAC 2510  
Db 2648 ATCCCGAAGCGCTGCATTCAAAAGTTCTCGCCTCTGAACCTTTCCGACCGCTCTCAC 2707  
Qy 2511 CAAAGTCTGGCTGGGCTCGGTTCTGTTGAAATCTTTTACGCGGCAATGGCCCTTA 2570  
Db 2708 CAGGTCTTTCGCTGGGCCCTTGGGTCCGTCGAAATTTTCTTCAGCAAGCACTGCCACTT 2767  
Qy 2571 TGGTATGCTACGAGGGCGCTCAAGTCTCTGGAGAGATTCGGTTATACATCAACACACC 2630  
Db 2768 TGGTACGGATACGGCGGGCTAAAAATTCCTGGAAAAGTTTCTTATATCAACTCCATC 2827  
Qy 2631 ATTTACCACATACTCTCTCCGCTTCTAGTCTATTGTATATTGCTGCTATCTGTCTG 2690  
Db 2828 GTTTATCCTGGACGTCCATTTCTCTGCTTACTGTACCTTGCCTGCGCATCTGCTG 2887  
Qy 2691 CTCACTGAAAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGTTTCATTGGC 2750  
Db 2888 CTCAGGGGAAGTTTATCACACAGAGCTTACCAATGTCGCCAGTATCTGTTTCATGGCA 2947  
Qy 2751 CTCCTCTTTCAATTTTCGCACTGGTATCCTTGAATGAGGTGAGTGGTGGTTCGAT 2810  
Db 2948 CTTTTCATCTGCTATCCGTCGCGGCTATCCTGGAAAATGAGGTGGGTTGGCCATC 3007  
Qy 2811 GACGAGTGGTGGAGAAAGAACAGTCTCTGGGTCATTGGAGGTATCTTGCACATCTGTTT 2870  
Db 3008 GACGACTGGTGGAGAACGAGCTTCTGGGTCTATCGGAGCGGTTTGGCGCATCTGTTC 3067  
Qy 2871 GCGCTCTTTCAGGGTCTTCTGAAGTGTTCGCCGTATCGACACAACTTCACCTGTGCACC 2930  
Db 3068 GCGGTGTTCCAGGGCTCTCTGAAGGTGTCGCCGCGCATCGACAGAGCTTCACCGTGACG 3127  
Qy 2931 TCAGAAGGTATAGCAAGAAGAGGAGCTTTCTGAGCTCTACATGTTCAAGTGACGACG 2990  
Db 3128 TCGAAGGCGGGGAGCGACGA---GGAGTCTCGGAGCTGTACAGGTTTCAAGTGACCAACC 3184  
Qy 2991 CTTCTCATCCTCCGACGACCATTTTGATCATTAACATGGTGTGGTGTGCTGCTGGCGACC 3050  
Db 3185 CTGCTGATACCCCGACCCAGCCGCTCCTCTGCTGTAACCTTCATCGGGGTGGTGGCGGGATC 3244  
Qy 3051 TCCTACGCCATCAACAGTGGTTACCAATCATGGGGCGGCTCTTTTGGGAAAGCTCTTCTTT 3110  
Db 3245 TCGAAGCGGATCAACAACGGGTACGAGTCGTGGGCGCCCTGTTCGGGAAGCTCTTCTTC 3304  
Qy 3111 GCCTTCTGGGTGATTTGTTACCTTATACCCATTCCTCAAGGTCCTTATGGGAGGCAAAAC 3170  
Db 3305 GCCTTCTGGGTGATGTGCTCCACTGTACCGTTCTCAAGGGTCTGGTGGGAGGAGAAC 3364  
Qy 3171 CGCACACGAGATTGTCTGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3230  
Db 3365 AGGACCGCGAGATGCTCATCTGTTGCTTCATCTGTTGCTTCATCTCTCTCTCTCTCTCTG 3424  
Qy 3231 TGGGTCTGTTGATGCCATTCACCTACCGCTCTGCTGGCCCAAAATATCCAAACCTGTGGC 3290  
Db 3425 TGGGTCCGCTGACCGGCTTCTCCGCAAGAGCAACGCGCGCTCTCTGGAGGAGTGTGGC 3484  
Qy 3291 ATCAACTGCTAGGAAAGTGGAG 3313  
Db 3485 CTGGACTGCAACTGAAAGTGGGG 3507

RESULT 5  
US-09-900-237-3  
; Sequence 3, Application US/09900237

Patent No. US20020120124A1		Query Match		36.4%; Score 1320.4; DB 10; Length 3776;	
GENERAL INFORMATION:		Best Local Similarity		64.5%; Pred. No. 0;	
APPLICANT: Allen, Stephen		Matches 2135; Conservative		0; Mismatches 1071; Indels 102; Gaps 7;	
TITLE OF INVENTION: Plant Cellulose Syntheses		QY 47		GCTCTGCGCGCCATGGACGGGACGCGGACGCGCCCTGAAGTCGGGAGGACGCGGGCGG 106	
FILE REFERENCE: BB1170 US CIP		DB 226		GCTCGTGTGTCATCGCGCGGATGGGAGCCAGCGCGGCGGAGCCCATGGACACGCGGAACGG 285	
CURRENT APPLICATION NUMBER: US/09/900,237		QY 107		GGACGTTGTCAGATCTGGCGCCGACGCGCTTGGCACACACGTTGGAGCGGACGCTCTTAC 166	
CURRENT FILING DATE: 2001-07-06		DB 286		CCAGGTTGTCAGATTTGGCGGACGACGCTGGGGCGCAACCCGACGCGGGAGCGGCTTCGT 345	
PRIOR APPLICATION NUMBER: 60/092,844		QY 167		CGCTCGGAGCTGCGGCTTCCGGTCTCGCGCTCGCGCCCTGCTAGCAGCAGCAGCGCAGGA 226	
PRIOR FILING DATE: 1998-07-14		DB 346		GGCTTGAACGAGTGGCGCTTCCCATCTGCCGGGACTGCTACGAGTACGAGGCGCGCGGA 405	
PRIOR APPLICATION NUMBER: PCT/US99/15871		QY 227		GGCACCCAGCGCTGCTCCAGTGCAGAGCAAGTCAAGCGCCACAGGAGGAGCGCCAGC 286	
PRIOR FILING DATE: 1999-07-13		DB 406		GGGACGACAGAACTGCCCCAGTGCAGACCCGCTTCAAGCGCCTCAAGGGGTGCGGCG 465	
PRIOR APPLICATION NUMBER: 09/720383		QY 287		GATCCGGGGGAGGAGGACACACACTGATGCCGATGATGGTGTAGTACTTCACTACCC 346	
PRIOR FILING DATE: 2000-12-21		DB 466		CGTCCCGGGGACGAGGAGGAGGAGCGGCGTCCGACGCTCGAGAAACGAGTTCAC----- 520	
NUMBER OF SEQ ID NOS: 33		QY 347		TGCACTGCGCACTGAGGACAGAGCAGAGATTGCTGACAGATGCGCAGCTGGAGCCAT 406	
SOFTWARE: Microsoft Office 97		DB 521		-----TGGAGGCGACAGCACACTCCAGTACCTCGCGGAGTCCATGCTCCACGCCACAT 576	
SEQ ID NO 3		QY 407		GAACACCGGGGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCT 466	
LENGTH: 3776		DB 577		GAGCTACGGCGCGGCGGCGGACCTCGACGCGGTGCGCGACCCATTTCCACCCCATCCCAA 636	
TYPE: DNA		QY 467		CTCAAGTATGACAGTGGAGATGCTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCT 526	
ORGANISM: Zea mays		DB 637		TGTTCCCTCCTCACCACAGGACAGATGGTTCGATGACATCCCGCGGACACGACGCGCT 696	
US-09-900-237-3		QY 527		GATGTACAGGAATCCCTGGAGCTTCCGCTGATCATCATGATGTCCTTACCGGGAA 586	
		DB 697		TGTCCTCGTTCGTGGTGGCGGGGAGAGAGATTACCCCTCT---CCCGTACGGGA 753	
		QY 587		CATCAGACAGTGTCTCGCTTCCCTATGTGAATTCATACCAAAATCCGTCAGGGAGTT 646	
		DB 754		TCCCAACTTCTGTGTCAACCGAGGTCTATGGACCCCTTCCAAAGGATCTCGCGCATATGG 813	
		QY 647		CTCGGCGAGTATGGGAATCTGCTGGAAGAGAGAGTTGATGGCTGGAATGAAGCA 706	
		DB 814		CTAC-----GGGAGCGTAGCATGGAAGGAGGAGGATGGAGCTGGGAAGCAGAAGCA 864	
		QY 707		GGACAAGGTTGCGATTCCCATGACTAATGGGACAAGCATTTGCTCCCTCTGAAGTCCGGC 766	
		DB 865		GGAGAGGATGCACACGACGAGGACGATGGCGGGCGGATGGTGTATGA----- 915	

QY 767	AGCTACTGACATCGATGCACT	826	AATACAACTGGAAGAGCGCTTTTACTGTAATGATGA
DB 916	-----TGCAGATCTACCACTAATGGATGA	939	-----TGCAGATCTACCACTAATGGATGA
QY 827	AACTCGCCAGCCCTCTATCTAGA	886	CCCCATTTGCTTCTCCAAATAAATCCCTACAG
DB 940	AGCTAGACAGCCATTTGCCAGA	999	CCCGCTTCTTCAAGCCAAATCAACCCCTATAG
QY 887	AATGGTCAATTTGCTCGGGTTC	946	CTCAAGCATCTTCTGCACCTACCGTCTCAACA
DB 1000	GATGATATAATAATTCGGGTA	1059	TTTGTGTTCTTCTTCCACACCGAGTGATGCA
QY 947	TCCTGTGCGTAATGCAATACCA	1006	SGCTTTATCTGTTATATGTCAGATTTGGTTGC
DB 1060	TCGGTGCCTGATGCAATTTGCT	1119	SGCTCATATCTGTGATCTGTAATTTGGTTGC
QY 1007	TTTATCCTGGATACTGGATCAG	1066	GAAGTGGTTTCCCAATCAACCGGAGACCTACCT
DB 1120	CATGCTTGGATTTCTTGACCA	1179	CAAGTGGTTTCCCTATCGAGAGGAAACCTATCT
QY 1067	TGATAGACTGGCTTTAAGTAT	1126	SAGAAGTGAACCGTCTCAGTTGGCTGCTTTGA
DB 1180	TGACCGGCTGAGTTTAAGGTT	1239	AGGAAGGCATCTTCTCACTCGCCCTGTTGA
QY 1127	CATATTTGTGACGACGACGAC	1186	TGAAGGAGCCACCTATCTCCTACTGCCAACACTGT
DB 1240	TTTCTTTGTGACGACGCTTGT	1299	TGAAGAACCTCCATTTGCTACTGCTAATACTGT
QY 1187	GCTATCCATTTCTGCTTTGAT	1246	CCGTGGACAAGGCTCTCTGCTATGATCTGATGA
DB 1300	TCATCTATCTTTCGGTGGAT	1359	CAGTTGATAAGGTTTCATGCTACGTTCTGATGA
QY 1247	CGAGCTTCAATGCTGACTTTT	1306	CATTGGCTGACACTTCAGATTTGCTAGGAATG
DB 1360	TGGTGTGCTGCTGCTGACATT	1419	CATTGCTGAAACATCTGAATTTGCAAGAATG
QY 1307	GGTACCATTTGTGAAGAATTA	1366	TGAACCCAGAGCTCCCGAGTTTACTTTTGCCA
DB 1420	GGTTCTTCTGCAAAAGATA	1479	TGAGCCCTGCTGCTCCAGAGTGGTACTTCCAAC
QY 1367	GAAATTTGATTACCTGAAAG	1426	TCCAGCCCTTCAATTTGTTAAAGACCCCGGCGAT
DB 1480	GAGATAGACTACTGAAAGAC	1539	TGGCCCAAACTTTGTTAGAGAACGGAGAGCAAT
QY 1427	GAGAGAGAAATTAAGAATTA	1486	TCAGATAAATGCCCTAGTTTCTAAGGCATTTGAA
DB 1540	GAGAGAGAGTATGAGGAATT	1599	TCAGAATCAATGCCCTTGGTTGCTAAGGCCCAAA
QY 1487	AGTCCCGGAGAGGATGGATG	1546	SAGATGGCACACCATGGCCAGGAAACATACCAG
DB 1600	GGTTCTCTGAGGAAGGATGG	1659	AGGATGGAACTCCATGGCCCGGAAATAATGTCCG
QY 1547	GGATCATCTCTGGAATGATTA	1606	TCCTTGGTCACAGTGGTGGCTTGTACTGTAGGG
DB 1660	TGATCATCTCTGGAATGATTA	1719	TCCTTGGTCAAAAGTGGTGGCCATGATGTGAAG
QY 1607	TAATGAGCTCCCCGTTTAGTT	1666	SGTCTGTGAAAAGCGTCTGCGGTTCCAGCACCA
DB 1720	AAATGAGCTGCTCGATTGGTT	1779	TTTCAAGAGAAAAACGGCCAGGCTACAACCATCA
QY 1667	CAAGAAGGCTGGTCCCATGAT	1726	TTGTTCTGCTCTCAGCTGTCTTACTTAATGGACA
DB 1780	CAAGAAGGCTGGTCTATGAA	1839	TGGTCCGAGTCTCTGCTGCTACTTAATGCTCTCC
QY 1727	ATACATCTGGAATCTTGATTC	1786	ACTACATCAACAAACAGAGGCTGTCCGAGAACG
DB 1840	TTATTTGCTGAATTTGATTC	1899	ACTATATCAATAATAGTAAGCTATAAAGGAAGC
QY 1787	TATGTGTTCTTAATGATTC	1846	TAGTCCGAGAGTCTGTTATGTGAGTTTCCCAAC
DB 1900	AATGTGTTTATGATGATTC	1959	TTGGAAGAAAGTTTGTCTATGTGCAAGTTTCTCTCA
QY 1847	AAGTTTGTATGGGATTTGATG	1906	ATCGATATGCAACAGGAACACTGTCTTTTGA



Db 371 GGGACGACAGAACTGCCCCAGTGCAGAGACTCGATACAGCGCCTCAAGGGCTGCCAAG 430  
Qy 287 GATCGCGGGAGAGCGACACTGATGCCGATGATGGTAGTACTTCAACTACCC 346  
Db 431 TGTGACCGGTGACGAGGAGGAGCGCGTGCATGACCTGGACAACGAGTTCAACTGGGA 490  
Qy 347 TGCATCTGCGACTGAGGACGACAGAGAGATGCTGCAGAGATGCGCAGCTGGCGCAT 406  
Db 491 CGGCATACATC-----GCAGTCTGGCGGAGTCCATGCTTACGGCCACAT 538  
Qy 407 GAACACCGGGGCGAGTGGCAATGTTGGCCACCCCAAGTATGACAGTGGCGAGATCGGCCT 466  
Db 539 GAGCTACGGCGTGGAGGTGA-----CCCTAATGGCGG 572  
Qy 467 CTCCAAGTATGACAGTGGAGAGATCCCTAGGGATAGTCCCTTCAGTCAACCAAGCA 526  
Db 573 CCACAAGTGTTCAG-----CTCAACCCCAATGTTCCACTCCTCACCACCGGGA 622  
Qy 527 GATGTCAGGAGAAATCCCTGGAGCTTCGCTGATCATCATCATGATGTCCTCCTACGGGAA 586  
Db 623 AATGGTGAATGACATCCACCGGAGCAGCAGCGCTGGTCCCTTCTCATGGGTGGTG 682  
Qy 587 CATCAGCAGAGCTGCTCGGTTTCCCTATGTAATCATATCACAATCCGTCAGGGAGTT 646  
Db 683 GGGAAAGAGGATACATCCCTTCTTATGCGGATCCAGCTTACCTGTGCAACCCAGGTC 742  
Qy 647 CTCGCGCAGTATGGGAATGTTGCTGGAAGAGAGAGTGTGATGGCTGGAAATGAAGCA 706  
Db 743 TATGACCCATCCCAAGATCTGTGATATGGTATGGTAGTGTGCTTGGGAAGCAAG 802  
Qy 707 GGACAAGGGTGGATTCCTATGACTAATGGGACAAGCATTTGCTCCCTCAGAGGTGGGC 766  
Db 803 GATGGAGAAATGGAGAGAGACACAAGAGAGATGCACGACGGGAAATGATGGTGGTG 862  
Qy 767 AGCTACTGACATGATCATCTACTGATACATACATGAGAGAGCTTTACTGATGATGA 826  
Db 863 TGATGATGGTGACATGCT-----GATCACCACATATGATGA 901  
Qy 827 AACTCGCAGCGCTCTATAGAAAAGTCCCAATGCTTCCTCCAAAATAATCCCTACAG 886  
Db 902 AGCAAGACAACACTGCCAGGAAATTCACATTCACATCAAGCCAGATTAATCCATATAG 961  
Qy 887 AATGGTCAATTTCTGGGTTGGTTGTTCTAAGCATCTTCTCGCATCCGCTACCGTCTCA 946  
Db 962 GATGATTAATCATTAATTCGGCTTGGTTTGGGTTCTTCTTCCACTACCGAGTGATGA 1021  
Qy 947 TCCTTGCGTAAATGCATACCACACTGCTGGCTTTATCTGTTATATGAGAGATTTGGTTGC 1006  
Db 1022 TCCGGTGAATGATGATTTGCTTTGGCTCATATCTGTTATCTGGAATCTGGTTGC 1081  
Qy 1007 TTTATCTGGATGATGATGATTCGCGAAGTGGTTTCCCAATCAACCGGGAGACCTACCT 1066  
Db 1082 CATGCTCTGGATTTGATCAATTCCTCAAAAGTGGTTCCTTATGAGAGAGACTTACCT 1141  
Qy 1067 TGATAGACTGCTTTAAGGTATGACGAGAGAGTGAACCTCTCAGTTGGCTGCTGTTGA 1126  
Db 1142 AGACCGGTGCTACTGAGGTTCGACAAGGAGGCCAGCCATCTCAACTGCTCCCAATGA 1201  
Qy 1127 CATATTTCTCAGTACAGTTCGACCCCTTGAAGGAGCCACCTATCGTCACTGCCAACACTG 1186  
Db 1202 TTTCTTTGCTAGTAGGTTGATCCCTTAAGGACCTCTTTGGTCACACAAATACTGT 1261  
Qy 1187 GCTATCCATTTCTGCTGTGATATATCCGTGGACAGGCTCTTGTGATGATGATGATGA 1246  
Db 1262 TCTATCTATCCTTTCCGGTGGATATCTCTTGAAGGATTTCTGCTATCTTTCTGATGA 1321  
Qy 1247 CGGAGCTTCAATCTGACTTTTGAAGCATTTGGCTGAGACTTCAGATTTGCTAGGAAATG 1306  
Db 1322 TGGTGTGCAATGCTACGTTTGAAGCATTTATCTGAAACATCTGAATTTGCAAGAAATG 1381  
Qy 1307 GGTACCAATTTGTGAAGAGTATGACATTTGAACCCAGAGCTCCCGAGTTTACTTTGGCA 1366  
Db 1382 GTTCCCTTCTGCAACCGGTACATATTTGAACCTCGCGCTCCAGAGTGGTACTTCCACA 1441

Qy 1367 GAAATTCATTACCTGGAAGA 1426  
Db 1442 GAAGATAGACTACTTGAAGA 1501  
Qy 1427 GAAGAGACAATATGAAGA 1486  
Db 1502 GAAGAGAGATATGAGGAAT 1561  
Qy 1487 AGTCCCGGAGGAAGATGGAT 1546  
Db 1562 AGTTCCCTGAAGAAGATGGAC 1621  
Qy 1547 GGATCATCTCGAATGATTCAG 1606  
Db 1622 TGATCATCTCGAATGATTCAG 1681  
Qy 1607 TAATGAGCTCCCGGTTTAGT 1666  
Db 1682 AATGAACCTGCCAGATTTGTT 1741  
Qy 1667 CAAGAGGCTGGTCCCATGAT 1726  
Db 1742 TAAGAAGCTGGTCTATGAA 1801  
Qy 1727 ATACATGTTGAATCTTGATT 1786  
Db 1802 ATATTTGTTAACTTGGATT 1861  
Qy 1787 TATGTCCTTCTAATGGATCC 1846  
Db 1862 AATGTTGTTTATGATGGACC 1921  
Qy 1847 AAGTTTCATGGATGATAG 1906  
Db 1922 AAGATTGATGGATGATCC 1981  
Qy 1907 TATTAACCTTGGGGCCCTGAC 1966  
Db 1982 TATCAACATGAAGGTTTGGAT 2041  
Qy 1967 TTTCAACAGACACTATCAT 2021  
Db 2042 ATTTAGAAGCGAGCATTTAT 2101  
Qy 2022 ----- 2059  
Db 2102 GACTTGCAACTGCTGGCCCAAG 2161  
Qy 2060 AAGCAAGTCAAGAAAAGGAGC 2113  
Db 2162 AAAGAAGACTACCAAAACCAAA 2221  
Qy 2114 CAGTCTCTCTCCAGTATTCAT 2173  
Db 2222 GAACCAATCCCTGCATATGCT 2275  
Qy 2174 TGATGATGAGAAATCAGTTCTC 2233  
Db 2276 TGAGAATCAAAAGGCCGATTT 2335  
Qy 2234 AGCAGCATTTGTCCTCCACT 2293  
Db 2336 TTCTGTTTTTGTATCATCCACA 2395  
Qy 2294 AGAATCTCTTTTGAAGAAGCT 2353  
Db 2396 TGTCTCTCTTTTGAAGAAGCT 2455  
Qy 2354 ATGGGAACTGAGATTTGGTGA 2413  
Db 2456 CTGGGAAACAGATTTGCTGGA 2515





1616 TGGAGGCTTCTGATCTGAAGGAACCACTTCCTCGCCTTCTTATGTTTCCAGAGAA 1675  
1646 GCGTCTGGGTTCCAGACACCAAGAGGCTGGTGCCATGAATGCCCTTGTGTGCTC 1705  
1676 AAGGCGTGGTCTTCAACACCAAGAGGCTGGTGCCATGAATGCTGTGGTGGGATC 1735  
1706 AGCTGTCTTACTATATGACAAATCATGTTGAATCTTCAATGTTGATCACTACATCAACAA 1765  
1736 TGCTGTCTCAAAATGCTCTTCAATGTTGAATGTTGATCACTATGTCATTA 1795  
1766 CAGCAAGGCTGTCGAGAGGATGTTGCTTCTTAATGATCAACACCTAGTCCGAGT 1825  
1796 CAGCAAGGCTGCCGAGAGGCTGCTTCTTATGAGGCCACAAACTGGGAAGAGGT 1855  
1826 CTGTTATGTCAGTTCACCAAGAGGTTGATGAGGATGATAGGATGATCATATGCAAA 1885  
1856 CTGTTATGTCAGTTCCTCAAGATTTGATGATGATGATGATGATGATGATGATGATG 1915  
1886 CAGGAACACTGCTCTTCTTATGATTAATCTTCAAGGCGCTTCAAGGATCAAGGACAGT 1945  
1916 CAGGACACAGTTTCTTGTATTAATCATGATGAGGCTAGATGATGATCAAGGTCCTGT 1975  
1946 TTATGTGGAACTGTTGTTTCAACAGAACAGCTATCTATGTTATGAGCCCCCAAT 2005  
1976 ATATGTGGGAGTGGATGTTTCAAGGAGCAAGCTTGTATGCTATAATCCTCCCAA 2035  
2006 TAAGCGGAAGAGCCAGGTTCTTGGCATCACTATGTTGGGGCAGAGAAGGCAAGCAA 2065  
2036 GGGTCCAAAGCGTCCAAAATGGTAAGCTGTGTTGGC----- 2076  
2066 GTCAAAGAAAAGGAGCTCAGATAAGAAAAGTCGAACAGCATGTGGACAGTCTGTTC 2125  
2077 --CGTGTGTTGAAGCCCAAGAGTATAAGGAGAGATGATGCAATGAGAGGCTGC 2134  
2126 AGTATTAATCTCGAAGACATAGAGAGGGTGTGAAGGTCGTGGGTTGATGATGAGAA 2185  
2135 AAGCCTAAA-----AGGGATGGATGATGACAA 2161  
2186 ATCAGTTCTCTGCTCAAAATGAGCTTAGAAGAGATTTGGCCAGTCAGAGCATTTGT 2245  
2162 AGAGGTGTGATGTCCTCAATGAATTTGAGAAGAAATTTGACAAATCCTCTATTTTGT 2221  
2246 TGCTCCACTCTGATGAATATGGTGTGTTCTCAGTCCCTCCACTCCAGAACTCTTTT 2305  
2222 GACTCTACCTTGATGAAGAGGGTGTGCTCCTCTTCAAGTCAGCTGCCCTGCT 2281  
2306 GAAAGAGCTATCCATGTCATAAGTTCGCTATGAGGACAGCTGAATGGGAACTGA 2365  
2282 TAAGAGCCCATTCATGATGATGCTGGATATGAAGATAAACTGAATGGGACTTGA 2341  
2366 GATTGGTTGGATCTGTCACAGAAAGATATTTACTGGATTCAAGATGCAGGC 2425  
2342 GCTTGGTTGGATCTATGATCTATCAGAGAAGATATCTACAGGTTTAAGATGCATG 2401  
2426 AAGAGGCTGGGTTCAAGTCTATGATGCCCAAGGCCAGCTTCAAGGATCTGCCCC 2485  
2402 CCGTGGGTGGAGTCCATTTATGTTATGCCAAAGAGAGCTGCAATCAAGGTTACTGCTCC 2461  
2486 CATCAATCTTTCAGATCTGTAACCAAGTCTCGGGTGGGCTCTCGGTTCTGTTGAAT 2545  
2462 TATCACTTGTAGATGCTCTCAACCAGGTTCTTGGTTGGGACTTGGTTCCATGAGAT 2521  
2546 TCTTTTACGCCGATTCGCCCTTATGATGGCT---ACGGAGGGCGCCCTCAAGTTTCT 2602  
2522 TTTCTTGTAGTACCATTGCCCTCTATGATGCTTCAAGGAAAAAGAGCTAAAGTGCT 2581  
2603 GGAGAGATTCGTTACATCAACACCACTTTACCCTAACCTCTCTCCCGCTTCTAGT 2662  
2582 TGAGAGATTTGCCATGCAAAACAACTGCTATCCATCCCAATCTCTAGTTGC 2641  
2663 CTATTTGATATGCTGCTATCTGCTCTCACTGGAAGTTTCATCATGCGCAGAGATTAG 2722

2642 CTACTGTATTTCTTCCAGCAGTT 2701  
2723 CAACCTTGCCAGTATCTGGTTG 2782  
2702 CACCTTGTGCTGGTTGTACTTT 2761  
2783 TGAGATGAGGTGGAGTGGTGT 2842  
2762 TGAGTTGAATGGAGTGGAGT 2821  
2843 CATTGGAGGTATCTTCACAT 2902  
2822 CATTGGTGTATCAGCTCA 2881  
2903 CGGTATCGACACCACTCAC 2962  
2882 TGGAAATTGACACCAATTTCA 2938  
2963 TGAGCTCTACATGTTCAAGTG 3022  
2939 AGAATTGTACACCTTTAAGTG 2998  
3023 TAACATGTTGGTGTGCTTGC 3082  
2999 TAACATTTGGTGTGTTGC 3058  
3083 GGGCGGCTCTTGGGAAGCT 3142  
3059 GGGACACTCTTGGAAAGCT 3118  
3143 CCTCAAGGCTTATGGGCAG 3202  
3119 CCTTAAAGTTGATGGTGC 3178  
3203 CCTCCTGCTCTATCTCTC 3262  
3179 GCTATGGCCCTATTTCTC 3238  
3263 CGCTGGCCCAATATCCAAAC 3315  
3239 TAAGGGACCTGATACCAAGCT 3291

RESULT 8  
US-09-838-539-1  
; Sequence 1, Application US/09  
; Patent No. US20020129401A1  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al  
; TITLE OF INVENTION: Plant Ce  
; FILE REFERENCE: 15621/03/US  
; CURRENT APPLICATION NUMBER:  
; CURRENT FILING DATE: 2001-0  
; PRIOR FILING DATE: 1996-10-2  
; PRIOR APPLICATION NUMBER: 08  
; PRIOR FILING DATE: 1997-10-2  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Window  
; SEQ ID NO 1  
; LENGTH: 3328  
; TYPE: DNA  
; ORGANISM: Artificial Sequen  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-09-838-539-1  
Query Match 33.5%  
Best Local Similarity 69.2%  
Matches 1710; Conservative  
QY 807 GACGCTTACTGAATGATGAA

se Synthase and Promoter  
/838,539  
987  
048  
sion 4.0  
gonucleotide  
ore 1215.6; DB 10; Length 3328;  
ed. No. 0;  
Mismatches 739; Indels 21; Gaps 3;  
CCAGGCTCTATCTAGAAAAAGTCCCATTCCTTCC 866





Db 2141 GATCTGAAGCAACAATGAAATTTGGAGAAGAAGTTTGACAAATCTCCGGTTTCGTTG 2200  
Qy 2248 CCTCAGCTCTGATGGAATATAGTGGTCTTCTCAGTCTCCTCACTCCAGAAATCTCTTTTGA 2307  
Db 2201 CCTCTGCTGTCTACAGACGGTGGAGTTCCCGTACGCAAGCCCGCATGTTTGTAA 2260  
Qy 2308 AAGAACTATCCATGTCTAAGTTGTGGCTATGAGGACAAGTCTGAATGGGGAACTGAGA 2367  
Db 2261 GAGAAGCCATTTCAAGTTATTAGCTGCGGGTACGAAGATAAAACCGATGGGGAAGAGA 2320  
Qy 2368 TTGTTGATCTATGAGTCTGTACAGAAGATATCTTACTGGATTCAGATGCACGCCAA 2427  
Db 2321 TCGGGTGAATTTAGATCGGTGACTGAAGATATCTTCGACGGGTTCAGATGCATTGCC 2380  
Qy 2428 GAGGCTGGGTTCACTATTTGATGCCCAAGCGCCAGCTTTCAGGGATCTGCCGCCA 2487  
Db 2381 ATGGATGGAGATCTGTGTACTGTATGCTAAGCGTGCAGCTTTTAAAGGATCTGCTCCTA 2440  
Qy 2488 TCAATCTTTACAGATCGTCTGAACCAAGTCTGCGGTGGGCTCTGGTTCGTTGAAATTC 2547  
Db 2441 TTAATCTGTCAATCGCTTCATCAAGTCTTACGTTGGGCTCTGGCTCTGTAGAGATTT 2500  
Qy 2548 TTTTCAGCGCGCAATGGCCCTTATGGTATGGTACGGAGGGCGCTCAAGTTCCTTGAGA 2607  
Db 2501 TCTTGACGACATTTCCGATATGGTATGGTATGGTGTAAATGGTTGGAGA 2560  
Qy 2608 GATTGCGTTTACATCAACACACCAATTTACCACATACTCTCCCGCTTCTAGTCTATT 2667  
Db 2561 GATTCTCTTACATCAACTCTGCTCTATCTCTGAGCTTCCACTTCCATTTGATGCTATT 2620  
Qy 2668 GTATATTGCCCTGCTATCTGCTGCTCACTGGAAGTTTCATCAATGCCAGAGATTAGCACT 2727  
Db 2621 GTTCTCTCCCGCGGTTTGTACTCACAGGAAATTCATCGTCCCTGAGATAGCAACT 2680  
Qy 2728 TGGCAGATCTGTTGCTGCTCTTCTTCAATTTTGGCCACTGGTATCTCTGAGA 2787  
Db 2681 ACGCAGGATATCTCTCATGCTCATGTTTCATATCCATAGCAGTAACCTGGAATCCTCGAAA 2740  
Qy 2788 TGAGTGGAGTGGTGTGGCAATGACGAGTGGTGGAGGAATGAACAGTCTTGGGTCAATG 2847  
Db 2741 TGCAATGGGAGGTGTGCGAATCGATGATTTGGTGGAGAAACGAGCAGTCTTTGGGTAATCG 2800  
Qy 2848 GAGGTATCTGCAATCTGTTTTCGCTTTCAGGGTCTTCTCAAGGTGCTTCCCGGTA 2907  
Db 2801 GAGGCGCTCTCCGATCTATTGCTCTGTTTCAAGTTTGGCTCAAAGTTCTAGCCCGAG 2860  
Qy 2908 TCGACACCAACTCTCACTCACTCAAGGCTAATGACGAAGAGGCGACTTTTGGTCTGAGC 2967  
Db 2861 TTAACACCAATTTACAGTCACTTCAAAAGC---AGCAGACGATGGAGCTTTCTCTCAGC 2917  
Qy 2968 TCTACATGTTCAAGTGGAGCGGCTTCTCATCCCTCCGACGACCATTTTGATCATTTAACA 3027  
Db 2918 TTTACATCTTCAAGTGGACAATTTGTTGATTTCTCCGACAAACACTTCTGATCATTTAACA 2977  
Qy 3028 TGGTTGGTGTGCTGCTGGCACTCTACGCCATCAACAGTGGTTACCAATCATGGGGC 3087  
Db 2978 TCAATTTGGAGTTATTGTCGGCTTCTGATGCCATTAGCAATGGCTATGACTCATGGGGAC 3037  
Qy 3088 CGCTCTTTGGGAAGCTCTTTTGCCTTCTGGGTGATGTTTCACTTTATACCATTCCTCA 3147  
Db 3038 CTCTCTTTGGGAGACTTTTCTCTGCTCTTTGGGTCAATTTGTTTCAATTTATACCATTCCTCA 3097  
Qy 3148 AGGCTTTATGGGAGGCAAAACCCACACGACGATGTCATGCTTGGGCTGCTCTCC 3207  
Db 3098 AGGGAATGCTTTGGGAAGCAACAAAATGCGCTACGATTTATTGGTGTCTGTTCTATTCTTC 3157  
Qy 3208 TCGCTTCTATCTTCTCTGCTGCTGGGTTCGTTGATGCCAAT 3250  
Db 3158 TAGCTTCGATCTTGACACTCTTTGGGGTCAAGATTAAACCGGTT 3200

US-09-938-842A-2530  
; Sequence 2530, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRIPT300-3  
; CURRENT APPLICATION NUMBER: US/09/938.842A  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2530  
; LENGTH: 3255  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-2530

Query Match 32.9%; Score 1191.4; DB 9; Length 3255;

Best Local Similarity 68.8%; Pred. No. 0;  
Matches 1687; Conservative 0; Mismatches 746; Indels 18; Gaps 3;

Qy 815 ACTGAATGATGAACTCGCCAGCCTCTACTAGAAAAGTCCCCATTGCTTCCTCCAAAAT 874  
Db 756 AATGATGATGAGGAGGAGCCGCAATGCTAGGAGATACCAATCAATCGACAAGT 815  
Qy 875 AAATCCCTACAGAAATGTCATTGTTCTGGGTTGGTGTGTTCTAAGAGATCTTCTCGACTA 934  
Db 816 AAATCCCTACCGATTTAAATGTCAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 875  
Qy 935 CGCTCTCAAAATCCGTCGCGTAATGCATACCCACTGTGGCTTTTATCTGTTATATGTA 994  
Db 876 CGTATCTTACCCCGTCAAGATGCATATGCTTTGGGCTTATCTGTTTATGTA 935  
Qy 995 GATTTGGTTGGTTTATCTCTGGATCTGTCAGTCTCCGAAAGTGGTTTCCAAATCAACCG 1054  
Db 936 GATATGTTGCTGTTTTCATGGTCTTCTGATCAGTCCCTAAATGTCACCTTATCGAGCG 995  
Qy 1055 GGAGACCTACCTGTAGACTGGCTTTAAGGTATGACCAGAGAGGTGAACCGTCTCAGTT 1114  
Db 996 AGAAAGCTACTTGGACCGACTCTCATTAAGATATGAGAAAGAAAGGAAACCGTCCGGACT 1055  
Qy 1115 GGCTGCTGTGACATATTTGTCAGTACACTCGACCCCTTGAAGGAGCCACCTATCGTCAAC 1174  
Db 1056 ATCCCTGTGGATGATTTGTTAGTACATGGATCCATTGAAAGAGCCCTCCGCTTATTAC 1115  
Qy 1175 TGCCAACTACTGCTATCCATTTCTGCTGTTGATTATCCCGTGGACAAGGTCTCTTGTA 1234  
Db 1116 TGCAAACTACTGCTCTTCTTATTTCTGCTGTTGATTATCTCGATAGAGTTCGTTGTTA 1175  
Qy 1235 TGATCTGATGACGAGCTTCAATGCTGACTTTTGACGCAATTTGGCTGAGACTCAGAGTT 1294  
Db 1176 CGTATCTGATGATGGTGGTCTGCTATGCTTACTTTCGAACTCTTTCTGAGACCGCTGAAT 1235  
Qy 1295 TGCTAGGAATGGTACCATTGTTGAAGAGTATGACTTGAACCCAGAGCTCCCGAGTT 1354  
Db 1236 CGCAAGAAATGGGTTCCTTCTGCAAGAAATATGTTATTGAGCCCTCGTCTCCGAAATG 1295  
Qy 1355 TTACTTTTGCAGAAAATGATTACCTGAAAGACAAGTCCAGCCCTTCAATTTGTTTAAAGA 1414  
Db 1296 GTATTCTGCCATAAAATGGACTACTTGAAGAAATAAGTTTCAATCCGCACTTTGTTAGGGA 1355  
Qy 1415 CCGCCGGGCAATGAAGAGAGATATGAAGATTTTAAATCAGGATAAATGCCCTAGTTTC 1474  
Db 1356 GCGCGGCAATGAAGAGAGATTTAAGAAATTTCAAAGATTTCAAAGTAAAGATCAATGCTTTAGTAGC 1415

QY 1475 TAAGGCAATGAACTCCCGAGGAGGATGATCATCAGATGGACACACCATGGCCAGG 1534  
D 1416 AACAGCAGAGAAAGTGCTGAGGATGGTTGACTATGCAAGACGGTACACCTGGCCCGG 1475  
QY 1535 AAACAATACCAAGGATCATCCTGGAATGATCAGGTTTCTTGGTGCACAGTGGTGGCT 1594  
D 1476 TAATAGTGTGCGAGATCATCCTGGCATGATCAGGTTCTTGGAGTGACGGTGTTCG 1535  
QY 1595 TGATACGTAGGAGTAATGAGTCCCGCGTTTGTAGTTATGTCTGTCGTAAGAGCGTCTGG 1654  
D 1536 TGATGTGGAACAAACAGAGTGGCTCGATTAGTTTACGTTTCTCGTGAGAAGACCGG 1595  
QY 1655 GTTCCAGCACACAAAGAGCTGGTCCATGAATGCCCTGTGTCGTTCTCAGCTCTCT 1714  
D 1596 ATTTGATCACATTAAGAGGCTGGAGCTATGAATTCCTGTAGAGGTTCTGGGGTCT 1655  
QY 1715 TACTAATGGCAATPACATGTTGAATCTTGTATCTGATCACTACATCAACAACAGCAAGC 1774  
D 1656 ATCAATGCTCTTACCTTGAATGTGATTTGTATCTACTACATCAACAATAGCAAGC 1715  
QY 1775 TGTCGAGAGAGCTATGCTTCTTAATGGATCCAAACCTAGGTCCCAAGTCTGTATGT 1834  
D 1716 TCTTAGAGAGCAATGTTTTCATGATGATGATCTCAGTCAAGGAAAGAAATCTGTATGT 1775  
QY 1835 GCAGTTCCTCCAAAGGTTTGTAGGATTTGATAGGAATGATCGATATGCAAAACAGGAACAC 1894  
D 1776 TCAGTTCCTCCAAAGGTTTGTAGGATTTGATAGGACAGTATCTCAATCGCAATGT 1835  
QY 1895 TGCTTTTGTGATTAATTAACCTTGAAGGCGCTTGACGCATTCAGGACCACTTATGTGGG 1954  
D 1836 TGTGTTCTTTGATTAATTAAGAGGTTTGGATGGCTTCAAGGGCTTATATAGTCTG 1895  
QY 1955 AACTGGTGTGTTTCAACAGAACAGCTATCTATGTTATGAGCCCGCCCAATTAAGCGGAA 2014  
D 1896 TACAGGTTGTGTTTCAAGAGGCAAGCGCTTACGGATTTGATGCAACCGAAGAGAA 1955  
QY 2015 GAAGCCAGG-----TTTCTTGGATCACHATGTTGGGGGCAAGAGGCAAG 2062  
D 1956 GGGCCACAGTAAGACATGCAATTTGCTGGCCAAATGCTCTCTATGTTTGGTTCAG 2015  
QY 2063 CAAGTCAAGAAAGAGGAGCTCAGA---TAAGAAAGAGTGAACAAGCATGTGGACAGTTC 2119  
D 2016 AAGAATCGTAAGCAAGACAGTGGCTGGGATAGAGAGAGAAATAGGAAGCGTC 2075  
QY 2120 TGTTCCAGTATCAATCTGAAGACATAGAGAGGAGTGTGAAGGTGCTGGGTTGATGA 2179  
D 2076 AAGCAGATCCAGCATTAAGAAATATCGAAGAGGCGCGTCACTTAAAGGTTCTAACGT 2135  
QY 2180 TGAGAAATCAGTTCTCATGTCTCAATGAGCTTAGAGAGAGATTTGGCCAGTCCAGCAGC 2239  
D 2136 AGACAGTCAACCGAGGCAATGCAATGAAGTTGGAGAGAAATTTGGCAGTCTCTCTGT 2195  
QY 2240 ATTTGTTCCCTCAGCTCTGATGAATATGTTGGTGTCTCTAGTCTCCACCTCCAGAAATC 2299  
D 2196 ATTTGTTGCTCTGCGGATGAGAAATGTTGGATGGCTAGAAAGCGAAGCCCGCTTG 2255  
QY 2300 TCTTTTGAAGAGAGCTATCACTGATTAAGTTGTGCTATGAGGACAACTGTAATGGG 2359  
D 2256 TCTGCTTAAAGAGAGGATCACTCACTATGTTGCGGATGAGAAATTAAGTAAATGGG 2315  
QY 2360 AACTGAGATTGGTGTGATCTATGATCTGTACAGAGAAATATCTTACTGGATTCAAGAT 2419  
D 2316 AAAAGAGATTGGTGTGATCTATGTTCTGTACCGAGATATCTTACGGGTTTAAAGT 2375  
QY 2420 GCAGGCAAGAGGCTGGGTTCACTCTATGTCATGCCCCAGCGCCAGCTTTCAAGGATC 2479  
D 2376 GCATTTCTCATGTTGGAGATCTGTTTATGTCACACCAAAAGTAGGCGCTTTCAAGGATC 2435  
QY 2480 TGCCCCCAATCAATCTTTTCAGATCTCTGAACCAAGTGTGCGGTGGCTCTCGTCTGT 2539  
D 2436 AGCTCCAAATCAATCTTTTCGATCGCTCTCCATCAAGTCTTTCGATGGGCGCTGGGTCTGT 2495

QY 2540 TGAATATCTTTTTCAGCGCGCAT  
D 2496 TGAGATTTTCTTGTAGTAGGCAT  
QY 2600 CTTGGAGAGATTTCCTTACATC  
D 2556 GCTTGGAGGTTGTCTTACAT  
QY 2660 AGTCTATTTATATTTGCTGCT  
D 2616 GCTTACTGTTCTCTCCCTGCT  
QY 2720 TAGCAACTTGGCCAGTATCTGG  
D 2676 TAGCAACTTGGAGTATCTCTG  
QY 2780 CTTTGGAGATGAGGTAGTGGT  
D 2736 TCTCGAGATGCAATGGGCAAA  
QY 2840 GGTCAATGGAGGTATCTTGCAG  
D 2796 GGTCAATGGAGGTGTTTGGG  
QY 2900 TSCCGGTATCGACACCACTTCA  
D 2856 TGCTGGTTCGACACTAACTTCA  
QY 2960 TGCTGAGCTCTACATGTTCAAG  
D 2913 CTCGACCTTACCTCTTCAAA  
QY 3020 CATTAAACATGTTGGTGTCTG  
D 2973 CATAACGTCATTCGAGTCATAG  
QY 3080 ATGGGGCGGCTCTTTGGGAAG  
D 3033 GTGGGACCGCTTTTCGGAAG  
QY 3140 ATTCCTCAAGGCTCTTATGGGCA  
D 3093 GTTCTTAAAGGTTTGGTGGG  
QY 3200 TGTCTCTCTCTCTATCTTCT  
D 3153 CATCTCTCTGCGCTCGATTCTTA

## RESULT 11

US-09-900-237-17  
; Sequence 17, Application US/099  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: plant Cell  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US  
; PRIOR FILING DATE: 2001-07-  
; PRIOR APPLICATION NUMBER: 60/0  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/7  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 17  
; LENGTH: 2890  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-900-237-17

## Syntheses

00, 237

4

15871



Query Match	32.6%;	Score 1183.4;	DB 10;	Length 2890;
Best Local Similarity	70.4%;	Pred. No. 0;		
Matches 1631;	Conservative	0;	Mismatches 671;	Indels 15; Caps
Qy	946	ATCCTGTGCGTAAATGCATACCCACACTGGCTCTTTATCTGTTATATATGTAGAGATTGGTTG	1005	
Db	12	ACCAAGTTAATGATGCATATGCGCTTGTGGTGCATCAGTCACTCTGTGAATATATGTTG	71	
Qy	1006	CTTTATCTCGGATACATGATCTCCGAAAGTGGTTTCCAAATCAACCCGGAGACCTACC	1065	
Db	72	CTGTATCTGGATAATGGATCAGTTTCCAAAATGGTACCCAAATACACGGAGAAACATACC	131	
Qy	1066	TTGATAGACTGGCTTTAAGSTATGACCGAGAGGTGAACCGCTCTCAGTTCGCTGCTGTTG	1125	
Db	132	TTGATCGTCTGTCACACAGTATGAAGAAAGGGAAGGCCATCTGAGTGTGCCAGTGTAG	191	
Qy	1126	ACATATTTGTCAGTACAGTCGACCCCTTTGAAGAGGCCACCTATCGTCACATGCCAACACTG	1185	
Db	192	ACGTCTTTGTCACTACTGTTGATCCCATGAAGGAACCTCCACTGATTACAGCAAACTG	251	
Qy	1186	TGCTATCCATTCTTGTGTTGATTATCCCGTGGACAAGTCTCTTGCTATGTATCTCATG	1245	
Db	252	TTCTATCTATCTTGTGTTGATTATCCAGTTGTATAAGTTTGCATGTATGTCTCAGATG	311	
Qy	1246	ACGGAGCTTCAATGCTGACTTTTGAAGTATGGCTTGAGACTTCAGAGTTTCTAGGAAT	1305	
Db	312	ATGGTCTGCTATGCTTACTTTTGAAGCACTGCTGAGACATCTGAATTTGCTAGAGAT	371	
Qy	1306	GGGTACCATTTTGAAGAAGTATGACATTTGAACCCAGAGCTCCCGAGTTTTCATTTTGC	1365	
Db	372	GGGTTCATTTTGTAGAAATACAAATATTAGCCCGCGGCACCAAGATGGTACTTTGGTC	431	
Qy	1366	AGAAATTTGATTACCTGAAGACAAGTCCAGCCTTCAATTTGTTAAAGACCCGCGGCCA	1425	
Db	432	AGAAGATGGACTATCTGAAATAAAGTACACCCAGCAATTTGTGAGGAAGAGAGACAA	491	
Qy	1426	TGAAGAGAGAATATGAAGAATTTAAATCAGGATAAATGCCCTAGTTTCTAAGGCATTGA	1485	
Db	492	TGAAGAGGGATTATGAAGAATTTAAGTGAGGATTAAACAGTTTGTGTGCACACGACAAA	551	
Qy	1486	AAGTCCCGAGGAAGATGGATCATGCAAGATGGCACACCATGCCAGGAACAATACCA	1545	
Db	552	AGGTCTCTGAGGATGGACCATGCAAGATGGGACTCTCTGGCCTGGAAATTAATGTGA	611	
Qy	1546	GGGATCATCTCGAATGATTCAGGTTTTCCTTTGGTCAAGTGGTGGCTTGATACTCAGG	1605	
Db	612	GGGATCATCTCGCATGATTCAGGTTCTCCTTTGGCAGGATGGTGTCTGATCTTCAAG	671	
Qy	1606	GTAAATGAGCTCCCCCTTTAGTTTATGCTCTCTGTAAGAGGCTCTGGGTTCCAGACC	1665	
Db	672	GAAATGAGCTACCCCTTTGGTCTAGTTTCTAGAGAAAGAGGCCAGGGTTTGATCACC	731	
Qy	1666	ACAAGAAGGCTGGTGGCATGAATGCCCTTGTTCGTCTCAGCTGTCCCTTACTAATGGAC	1725	
Db	732	ACAAAAGGCTGGTGCATGAATGCTCTGGTACGGCTTCAGCAATTAATCAATGCAC	791	
Qy	1726	AATACATGTTGAATCTTGATTTGTGATCACTACATCAACACAGCAAGGCTCTCCGAGAAG	1785	
Db	792	CCATCTCTCTGAATGTGATTTGTGATCACTACATTAACAATAGCAAGGCACTTAAAGAAG	851	
Qy	1786	CTATGCTCTTCTTAATGGATCAAAACCTAGGTCGCAAGTCTGTTATGTGTCAGTTCCAC	1845	
Db	852	CTATGCTTTTATGATGGATCTCACTAGGAAAGAGTTTGCTATGTGCAATTTCTCTC	911	
Qy	1846	AAAGGTTTGATGGGATTTGATAGGAATGATCGATATGCAAAACAGGAACACTCTCTTTTG	1905	
Db	912	AGCGATTTGATGGAATTTGATAGACATGATAGATATTCAACACAGAAATCTGTATTTTTCG	971	
Qy	1906	ATATTAACTTGAGGGCCCTTGACGCATTCAGGACCAGTTTATGTGGGAAGTCTGGTGTG	1965	
Db	972	ATATTAACTGAAGGATTTGGATGGGATACAAGTCCCAATATATGTGCGGAAGTGGATGTG	1031	
Qy	1966	TTTTCAACAGAACAGCTATCTATGTTATGAGCCGCCCAATTAAGGCGAA--GAAGC	2019	

Db	1032	TTTTTCAAGAGGTTACGCACATTTATGGGATATGATGACCTTGCACAAGAAAGAAACCCACGACGA	1091
Qy	2020	CAGGTTTCTTGGCATCAGTATGTGGGGGCAAGAAGCAAGCAAGTCAAAAGAAAGGA	2079
Db	1092	AAACTTGTAACCTGTTGCCCAAACTGGTGCCTATGTTGGCTCTAGAAGAAAAGA	1151
Qy	2080	GCTCAGATAAGAAAAAGTCGAACAAGCATGTGGACAGTTCTTCCAGTATTCAATCTCG	2139
Db	1152	ATGCCATAGTAAAGAGGAGAAAAGAGGAGGTTGAAGCACAAGTCGAAGCATCAAGACGAGA	1211
Qy	2140	AAGACATAGAGGAGGTTGTTGAAG-----GTGCTGGGTTTGATGATCAGAAAATCAGTTC	2193
Db	1212	TACATGCACCTTGAATAATATGAGCGGGGGAATGAAGGAAACCAACATCGAAGACATCCA	1271
Qy	2194	TCATGTCTCAAAATGAGCTTAGAGAAGAGATTTGGCCAGTCAGCAGCATTTGTTCCTCCA	2253
Db	1272	ATCTGACTCAAAACAAAGTTTGGAGAAGAGTTTGGACAGTCTCCAGTATTTGTAGCCTCCA	1331
Qy	2254	CTCTGATGGAATATGTTGGTGTTCCTCAGTCTCCACATCCACAGAATCTCTTTTGAAGAAG	2313
Db	1332	CACTTTGGATGATGGTGGAGTTCCACATGGCGTGAGTCTCGATCACTTTTAAAGAAG	1391
Qy	2314	CTATCCATGTCAATAAGTTGTGGCTATGAGGACAAGCTGAATGGGGAACACTGAGATTGGTT	2373
Db	1392	CCATCCAGGTCATCAGTTGTGGTTATGAAGACAAACAGAAATGGGAAAGATTGGGT	1451
Qy	2374	GGATCTATGATCTGCACAGAAGATATTCTTACTGGGATTTCAAGATGACGCGAAGAGCT	2433
Db	1452	GGATATATGTTCTGTGCACAGAGGATATCTTACTGGGATTTAAATGCAATTTGCCATGGTT	1511
Qy	2434	GGCGTTCAGTCTATTGCATGCCCAAGCGCCAGCTTTTCAAGGGAATGTCGCCCATCAATC	2493
Db	1512	GGCGGCTGTGTATTCGATCTCCTAAGCGGCTGCAATTAAGGGGTCTCGCGCTATCAACC	1571
Qy	2494	TTTCAGATCTGTGAACCAAGTCTCGGTGGGCTCTCGGTTCTGTTGAAATCTTTTCA	2553
Db	1572	TTTTCAGATCTGTGCACCAAGTTCTTCGGTGGGCTCTTGGGTCTGTGAGATTTTTTTCA	1631
Qy	2554	GCGGCAATCCGCTTATGATGATGGCTACGGAGGGGCTCAAGTTCTCTGGAGAGATTCG	2613
Db	1632	GCAGACATGTCCAATCTGGTATGGCTATGTGGTGGATTTGAATTTGTTGAACGATTTT	1691
Qy	2614	CTTACATCAACACACCAATTTACCCACTAACCTCTCTCCCGCTTCTAGTCTATTGTATAT	2673
Db	1692	CCTACATTAACCTCGGTCGATATACCTCCGGACTTCCCTCCCATTTGCTTACTGTACTC	1751
Qy	2674	TGCGTGTATCTGTCTGCTACATGGAAGTTTCATCATGCGACAGATTAGCACTTGGCCA	2733
Db	1752	TACCAGGCATATGCGTTCTGACTGGAAAATTTATGTACCCGAGATTAGCACTATGGCA	1811
Qy	2734	GTATCTGGTTCATTGGCGCTCTCTTTCAATTTTCGCCACTGGTATCTCTGAGATGAGGT	2793
Db	1812	GTCTTGTGTTTCATGGCCCTCTTCATATCCATTCAGCAACTGGCATCTTGAGATGCAAT	1871
Qy	2794	GGATGGTGTGGCATATGACAGTGGTGGAGGAATGAACAGTCTCTGGGTCAATTCAGAGTA	2853
Db	1872	GGGCGGTGTATGACATAGACGACTGGTGGAGGAACAGATTTTGGGTGATTCGGAGGTG	1931
Qy	2854	TCTCTGCACATCTGTTGGCGTCTTTCCAGGGTCTCTGAAAGTGGTTCGCCGTATFCGACA	2913
Db	1932	TTTCTCCCATCTATTTCGCCCTATTTCAGGGTTTACTGAAGGCTTGGCTGGGTGAACA	1991
Qy	2914	CCAACTTCACTGTCACTCAAGGCTTAATGACGAAGAAGCGGACTTTGCTGAGCTCTACA	2973
Db	1992	CAAACTTCACTGTGACCTCAAAAGC---AGCAGATGATGGAAATTTCTCAGAACTTTACA	2048
Qy	2974	TGTTCAAGTGGACGAGCTTCTCATCCCTCCGAGACCAATTTTGATCATTAACATGGTGTG	3033
Db	2049	TATTCAAGTGACATCACTCTTGATCCCTCCATGACTTTTACTTTATCATGAATATTGTGCG	2108
Qy	3034	GTGCTGTTGGCACCTCTCAGGCCATCAACAGTGGTTTACCAATCATATGGGGCGGCTCT	3093

Db	2109	GGGTGGTGTGCGGATCTCAGATGCCATCAACAATGGTTATGACTCATCGGGACCTCTCT	2168
QY	3094	TTGGGAAGCTCTCTTTTGGCTTCTGGGTGATGTTCACCTATACCCATTCTCAAGGTC	3153
Db	2169	TTGGTAGATTGTCTTTGGCAATTGTGGGTGATCCTCCATCTCTACCCCTCTTTGAAGGGT	2228
QY	3154	TTATGGCAGGCCAAAACCCACACCGACGATTGTGCATCGTCTGGGCTCTCTCTCGCTT	3213
Db	2229	TGCTTGGAAACAAGATAGATGCCAACCATATATTGGTTGGTCAATCCTCTCGCCT	2388
QY	3214	CTATCTCTCCTTGCTGTGGTTCGTGTGATGCCATT	3250
Db	2289	CCATCTTGACTCTCATGTCTGGGTGAGAATTAACCCGCT	2325

```

RESULT 12
US-09-900-237-23
; Sequence 23, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Staph Cellulose Syntheses
; FILE REFERENCE: Bbl170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 2663
; TYPE: DNA
; ORGANISM: picramnia pentandra
US-09-900-237-23

```

Query Match 29.5%; Score 1069.6; DB 10; Length 2663;  
Best Local Similarity 69.7%; Pred. NO. 5.1e-302;  
Matches 1514; Conservative 0; Mismatches 634; Indels 24; Gaps 4;

Qy	1094	AGAAGGTGAACCGCTCTCAGTGTGGCTGCTCTTGACATATTTGTGAGTACAGTCGACCCGCTT	1153
Db	11	AGAGGGGAAGCCATCTGAGCTAGCTGGCTTAGACATATTTGTGAGTACGGTGGATTCCTAT	70
Qy	1154	GAAGGAGCCACCTATCGTCTCAGTCGCCAACACTGTGCTATCCATTCTTGCTGTTGATTATCC	1213
Db	71	GAAGAACCTCCACTTATCACTGCAAACTACTGTGATTCACCTCCATTCGAGTTGATTATCC	130
Qy	1214	CGTGGACAAGTCTCTTGTTATGTATCTCATCAGCGAGGCTTCAATGCTGACTTTTGAGCG	1273
Db	131	AGTTGATAAAGTTACATGCTAGCTGCAGATGATGGTGTGCCATGCTTACTTTTGAAGC	190
Qy	1274	ATTGGCTGAGACTTCAGAGTTTGCTAGGAAATGGGTACCATTTGTGAAGAAAGTAGCAT	1333
Db	191	ACTTCTGAAACATCTGAATTTGCACGAAATGGGTCCCTTCTCTGAAGAAGTTAGCAT	250
Qy	1334	TGAACCCAGAGCTCCGAGTTTACTTTTGGCCAGAAAATTTGATTACCTGAAAGACAAGT	1393
Db	251	TGAGCCTCGAGCACCAAGATGGTATTCTCTCAGAAAGATGGACTATTTTGAAGAACAAAGT	310
Qy	1394	CCAGCCTTCAATTTGTTAAAGACCGCGGCCATGAAGAGAGAAATGAAGAATTTAAAT	1453
Db	311	ACACCCATCAATTTGTTAGGGAAGAGCGTCTATGAAGAGAGAAATATGAAGTATTTCAAAGT	370
Qy	1454	CAGGATAAATGCCCTAGTTTCTTAAGGCATTGAAAGTCCCGAGGAAGATGGATCATGCA	1513
Db	371	TCGGATAAATGGTTTGGTTGCCATGGCACAAGAAGTTCCCGAGGATGGTTGGACGATGCA	430
Qy	1514	AGATGGCACACCATGGCCACAGGAACAATACCAGGGATCATCCTCGAATGATTCAGGTTTT	1573

QY 2639 ACTAACTCTCTCCGCTCTAGTCTATTGTATATTGCTGCTGCTACTGCTGCTACTGG 2698  
Db 1565 TTGAGCTTCATTCCCTTTGCTTTTACTGACACTGCCAGCAATTTGCTTCTTACTGG 1624  
QY 2699 AAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTCGCTCTTCC 2758  
Db 1625 GAAATTCATTTGCTTGGATTAGCAACTATGCAAGCACTTGTATTATGCTCTCTCAT 1684  
QY 2759 TTCAATTTTGGCACTGGTATCCCTTGAGATGAGTGGATGGTGGTGGCAATTCAGAGAG 2818  
Db 1685 ATTATTGCTGCAACAGCACTTTTGGATGCAATGGGTGGTGGTGGTGGTGGTGGTGGT 1744  
QY 2819 GTGAGGAATGAACAGTCTTGGTCAATGGAGGTATCTCTGACATCTCTTTGGCGCTCT 2878  
Db 1745 GTGAGGAATGACCAATTTGGGTCAATGGAGGTATCTCTGACATCTCTTTGGCTCTCT 1804  
QY 2879 TCAGGCTCTTCTGAAGTGGTTCGCGGTATCGACACCAACTTCACCTGTCCACTCAAGGC 2938  
Db 1805 CCAAGTTTACTCAAGTTTCTGGCTGCTGTGTAACACAAACTTCACCTGTCCACTCAAGGC 1864  
QY 2939 TAATGAGGAAGAGGCGACTTTGCTGAGCTCTACATGTTTCAAGTGGACAGCGTCTCAT 2998  
Db 1865 ---TGTGATGAAGGGGATTTCTCAGAGCTTTACCTCTTCAAGTGGACAACTTATTAT 1921  
QY 2999 CCTCCGACGACCATTTTGATCAATTAACATGTTGTGTGCTGTGGCACCTCTCTACGC 3058  
Db 1922 TCCTCCACNACGTTGTTGATCAATAATATTGTTGGGTGGTGGTGGTGGTGGTGGTGG 1981  
QY 3059 CATCAACAGTGGTTACCAATCATGGGGCGCTCTTTGGGAAGCTCTCTTTGGCTTCG 3118  
Db 1982 CATCAATATGTTATGATTCATGGGCTCTTTGTTGCTGGCTATTTTTCGCAATTCG 2041  
QY 3119 GGTGATGTTCACTTATACCATCTCTCAAGGCTTTATGGGCGAGGCAAAACCGCACCC 3178  
Db 2042 GGTCAATGTGCCACTCTATCTCTTCTCAAGGTTTGTCTGGGAACAAGATCGCACTCC 2101  
QY 3179 GAGGATGTCATGCTGGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3238  
Db 2102 AACCATTTGTGGTCTGGTCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2161  
QY 3239 TGTGTATCCATT 3250  
Db 2162 AATAAATCCATT 2173

RESULT 13  
US-09-900-237-15  
; Sequence 15, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; PRIOR FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 15  
; LENGTH: 2125  
; TYPE: DNA  
; ORGANISM: Glycine max  
US-09-900-237-15  
Query Match 26.7%; Score 966.8; DB 10; Length 2125;  
Best Local Similarity 71.6%; Pred. No. 5.2e-272;  
Matches 1316; Conservative 0; Mismatches 507; Indels 15; Gaps 3;

QY 1474 CTAAGGCATTGAAAGTCCCGAGGAGGATGGATCATCAAGATGGCACACCATGGCCAG 1533  
Db 2 CCAAGATCGCAAGAGATGCGAGGAGAGGTTGGACAATGAGGATGGAAATCTCTTGGCGTG 61  
QY 1534 GAAACAATACACGAGGATCATCTGCAATGATTACAGTTTCTCTGGTGCACAGTGGTGCC 1593  
Db 62 GAAATATCTTAGGATCATCCGGAATGATTCAGGTGTTTTTAGTTCATAGTGGGGGGC 121  
QY 1594 TTGATPACTGAGGTTAATGAGTCCCGCTTTAGTTTATGTTCTCTGTAAGAGCGTCTG 1653  
Db 122 TGGATACAGATGGAATGAGTCCCTAGACTTGTATTATGTTCTCTGTAAGAGCGACCCAG 181  
QY 1654 GGTTCACACACACAAAGAGGCTGGTCCATGAATGCCCTTGTCTGCTGCTCAGCTGTCC 1713  
Db 182 GCTTCCACATCATCAAGAAGGCTGGAGCTATGAATGCTTTGATTCGAGTTTCTGCTGTCT 241  
QY 1714 TTACTAATGGCAATPACATGTTGAATCTTGAATGTTGATCACTACATCAACACAGCAAG 1773  
Db 242 TGACCAATGGTGCATATCTTCTGAATGGGATTCGATCACTATTTCATTAATAGCAAG 301  
QY 1774 CTGTCGAGAGAGCTATGTTCTCTTAATGGATCCAAAGCTAGTCCGCGCAAGTCTGTTATG 1833  
Db 302 CCTCAAGAAGAGGCTGTTTCTATGATGGATCCTGTTCTTGGAAAGAGACATGCTATG 361  
QY 1834 TGCAGTTCCCAACAAAGGTTTGTATGGATTGATAGGAATGATCGATGCAACACAGGACA 1893  
Db 362 TTCAATTTCCCTCAGAGATTTGACGCAATGACTTGCAGATCGATATGCCAATCGCAATA 421  
QY 1894 CTGCTTTTTCATATTAACCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 1953  
Db 422 TTGTTCTTTCATATCAACATGAAGGCTCAGGATGGTGTTCAGGGCCCACTCTATGTTG 481  
QY 1954 GAATGTTGTTGTTTTTCAACAGACAGCTATCTATGTTTATGAGCCCAATTAAGGCGA 2013  
Db 482 GAATGTTGTTGTTTTTCAATAGGCAAGCTTGTATGTTTATGATCTCTGTTTGAAGG 541  
QY 2014 A-----GAAGCCAGGTTTCTTGCACTACTATGTGGGGC--AAGAAGAGGCAAGCA 2064  
Db 542 AAGATTGGAACCTAACATTAATGTAAGAGTTGTTGGGTTCTAGANAAGAGGAAAGG 601  
QY 2065 AGTCAAGAAAGAGGCTCAGATAAGAAAAAGTCGAACAGCATGTGGACAGTCTCTGTTTC 2124  
Db 602 GTGGCAATAAGAAGTACAGTGACAAGAGAGGCGATGGGAAGAACTGAATCCACTGTAC 661  
QY 2125 CAGTATTCAATCTCGAAGACATAGAGGAGGCTGTGAAGGTTGCTGGGTTTTCATGATGAGA 2184  
Db 662 CCATATTTAATATGGAAGACATAGAGGAGGTTGTAAGGT-----TATGATGATGAAA 715  
QY 2185 AATCAGTTCTCATGTCTCAAAATGAGCTTAGAGAAGAGATTTGGCCAGTCAGCAGCATTTG 2244  
Db 716 GGACACTACTTATGTCTCAAAAGAGCTTGGAGAAGCGTTTGGTCAGTCTCCAGTTTGA 775  
QY 2245 TTGCCCTCACCTCTGATGGAATATGTTGTTTCTCTAGTCTCTCACTCCAGAAATCTCTTT 2304  
Db 776 TTGCTGCCACTTTCATGGAGCAGGCTGCAATTCACCTTCAACGAAACCTGCAACTCTTC 835  
QY 2305 TGAAGAAGCTATCCATGCTCATAGTTGCTGCTATGAGCAAGCTCTCAATGGGGAAGCTG 2364  
Db 836 TTAAGGAAGCAATCCATGTTTACAGCTGTGTTAGGAAGACAAGACAATGGGGCAAG 895  
QY 2365 AGATTGGTTGGATCTATGGATCTGTCAAGAAAGATATTCTTACTGGATTCAAGATGCACG 2424  
Db 896 AGATTGGATGATCTATGGCTCTGTGACAGAAAGATATCTTGCACGTGGGTTCAAGATGATG 955  
QY 2425 CAAGAGGCTGGGTTCACTCTATTGCAATGCCAAGCGCCAGCTTTCAAGGATCTCGCC 2484  
Db 956 CTGCTGGTTGGATTTCCATCTATTGCAATGCCAAGCTTCAAGGATTTAAGGTTTCTGCTC 1015  
QY 2485 CCATCAATCTTTCAGATGCTGTAACCAAGTCTGCGGTGGGCTCTCGGTTCTCTTGA 2544  
Db 1016 CTATCAATCTTTCAGATGCTCAATAGGTTCTCGGTGGGCTTGGGTTCAATTGAGA 1075

```
QY 2545 TTCTTTTACGGCGCATTTGCCCTTATGGTATGGCTACGGAGGGCGCTCAAGTTCTGTG 2604
Db 1076 TCTTTCTAAGCAGGCATTTGCCCTTGGGTATGGCTACAATGGGAAGTTGAAGCCCTCTGA 1135
QY 2605 AGAGATTGCGCTTACATCAACACACACCATTTACCCACTAACTCTCTCCCGCTTCTAGTCT 2664
Db 1136 TGAGGCTTGGTTATATTAACACCATTTGCTACCCGTTTACCTCAATCCCATGATTGCTT 1195
QY 2665 ATTGTATATTGCTGCTATCTGCTCTGCTCACTGGAAGTTTCATCATGCCAGAGATTAGCA 2724
Db 1196 ACTGTAGCTTCTCGCATTTTGTCTCTCTCACAATAAATTTATTTCTTGAGATAGCA 1255
QY 2725 ACTTGCCAGTATCTGGTTCATTTGGCTCTTCTTCAATTTTCGCCACTGGTATCCCTTG 2784
Db 1256 ACTTGCCAGTATGTTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1315
QY 2785 AGATGAGGTGAGTGTGTTGGCATTTGACGAGTGGTGGAGGAATGAACACTTCTGGGTCA 2844
Db 1316 AGCTTAGGTGGAGTGGGTCAGTATAGAAGACTGGTGGAGAAATGAACAGTTCTGGGTTA 1375
QY 2845 TTGGAGGTATCTGTCACATCTGTTTGGCTCTTTCAGGCTCTTCTGAAGGTGCTTGGCG 2904
Db 1376 TCGGTGGGACATCTGCGCATCTTCTGCTGTGTCCAGGGGCTTCTAAAGTCTTGGCTG 1435
QY 2905 GTATGACACCAACTTCATCTGTCACCTCAAGGCTTAATGACGAAGGCGACTTTGCTG 2964
Db 1436 GGATGATACAAAATTTTACTGTTTACATCGAAGGCATCGGACGAGGATGGGACTTTGCCG 1495
QY 2965 AGCTTACATGTTCAAGTGACGACGCTTCTCATCTCCCTCGACGACCATTTTGCATCTTA 3024
Db 1496 AGCTTTATGTTTAAATGGACATCACTTCTCATCTCCCTCTCAACAGTCTTATTGTGA 1555
QY 3025 ACATGTTGGTGTCTGTTGGTGGACCTCTTACGCCATCAACAGTGGTTACCAATCATGGG 3084
Db 1556 ATTTGTTGGGATGTTGGCTGGTGTATCTATGCCATAAACAGTGGTTACCACTCTTGGG 1615
QY 3085 GGCGGCTCTTTGGGAAGCTCTCTTTGGCTCTTGGGTGATGTTGCTATTAACCCATTC 3144
Db 1616 GTCCACTATTGGCAAGCTGTTCTTTGCTATCTGGGTCAATGGCCATCTATACCCATCT 1675
QY 3145 TCAAGGCTCTTATGGCAGCAAAACCGCACACGAGATGTCATCGTCTGGGCTGTCC 3204
Db 1676 TGAAGGCTCTTTGGGAGCAAAATCGTACCCCAACCATTTGTTATGTTGGTCTCGGTC 1735
QY 3205 TCTCTGCTTCTATCTCTCTGCTGCTGGGTTGCTGTGATCCATTCATCCCGTCTCG 3264
Db 1736 TTCTTCTCAATATTCTCTTCTGCTGCTGGGTGAGGATTGATCCCTTCACCTCTGATCCA 1795
QY 3265 CTGGCCCAATATCCAAACCTGTGGCATCAACTGCTAG 3302
Db 1796 ACAATTAACCAATGGTCAATGTGGCATCAACTGTTAG 1833
```

```
RESULT 14
US-09-900-237-27
; Sequence 27, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Syntheses
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1998-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 27
; LENGTH: 1560
```

```
; TYPE: DNA
; ORGANISM: Glycine max
; US-09-900-237-27
Query Match 22.7%
Best Local Similarity 78.6%
Matches 1010; Conservative
QY 1093 GAGAAGGTGAACCGTCTCTCAGTT 1152
Db 8 GGGAGGGGAACCATCACAGCT 67
QY 1153 TGAAGGAGCCACTATCGTCACT 1212
Db 68 TAAAAGAACCCCGCTTGTGAC 127
QY 1213 CCCTGGGCAAGGCTCTCTTGTCT 1272
Db 128 CAGTGGGATAAGGCTCCCTGTTA 187
QY 1273 CATTGGCTGAGACTTTCAGAGTT 1332
Db 188 CTCTGGCTGAGACATCAGAAT 247
QY 1333 TTGAACCCAGAGCTCCCGAGTT 1392
Db 248 TCGAACCTCGGGCACCTGAGTG 307
QY 1393 TCCAGGCTTCATTTGTTAAAGT 1452
Db 308 TTCAACCATCATTTGTTCAAAG 367
QY 1453 TCAGGATAAATGCCCTAGTTTC 1512
Db 368 TTGCGCATCAATGACACTTGTTC 427
QY 1513 AAGATGGCACACCATGCCGAGG 1572
Db 428 AAGATGGTACGCCATGGCCCTGG 487
QY 1573 TCCCTTGGTCACAGTGGTGGCCCT 1632
Db 488 TCTTGGGCCAAAGTGGAGGACT 547
QY 1633 TGTCTCGTGAAGAGCGTCCCTGG 1692
Db 548 TTTCTCGTGAAGAGCGTCCAGG 607
QY 1693 TTGTTCTGTCCTCAGCTGTCTCT 1752
Db 608 TTGTTCCGAGTGTCCAGCAGTCT 667
QY 1753 ACTACATCAACAACAGCAAGG 1812
Db 668 ACTACATAACAACAGCTAAAG 727
QY 1813 TAGGTCCGCAAGTCTGTTATGT 1872
Db 728 TTGGGAAATGTTTGTCTATGT 787
QY 1873 ATCGATATGCAACAGCAAGCA 1932
Db 788 ATCGATATGCAATCGCAATAC 847
QY 1933 TTCAAGGACCGAGTTTATGTGG 1992
Db 848 TTCAAGGCTCTGTTTATGTGG 907
QY 1993 ATGAGCCCCCAATTAAGGGCA 2046
Db 908 ACGAAGCTCTCTATTAACCCAA 967
QY 2047 GCAA---GAAGAAAGCAAGCA 2103
```

ore 821.4; DB 10; Length 1560;  
ad. No. 1.4e-229;  
Mismatches 266; Indels 9; Gaps 2;

```
Db 968 GTAACCGAAAGAGAGATCAAAATCTAGCAAGAAAGGCTCAGACAGAAAAAATCTAGCA 1027
Qy 2104 AGCATGTGGACAGTCTGTTCAGATATCAATCTCAAGACATAGAGAGGGGTGTGAAG 2163
Db 1028 AGATGTGACCAACTGTGCCATCTTTAGCTTGAGATATAGAGAGGGGTGGAAG 1087
Qy 2164 GTGCTGGTGTGATGAGAAATCAGTTCTCATGTCTCAATGTCTAGAGCTTTAGAGAAGAT 2223
Db 1088 GTGCTGATTTGATGATGAGAAATCACTATTATGTCACAAAATGAGCCTCGAGAAAAGGT 1147
Qy 2224 TTGGCCAGTCAGACAGATTTGTTGCCCTCCACTCTGATGGAATATGTGCTCAGT 2283
Db 1148 TTGTCAGTCTGCTCTTTGTTGCCCTCTACACTGAGAGAAATGTGCGGTTCCTCAGT 1207
Qy 2284 CTTCCACTCCAGAACTCTTTTGAAGAAGCTATCCATGTCATAGTTGTGGCTATGAGG 2343
Db 1208 CTGCAACTCCAGAACTCTTCTTAAGGAAGCTATTTCATGTTATCAGTTTCTGTGTACGAGG 1267
Qy 2344 ARAAGTCTGAATGGGNACTGAGAT 2368
Db 1268 ATRAATCAGAATGGGGAAGTGAGGT 1292
```

## RESULT 15

US-09-900-237-1

; Sequence 1, Application US/09000237

; Patent No. US20020120124A1

; GENERAL INFORMATION:

; APPLICANT: Allen, Stephen

; TITLE OF INVENTION: Plant Cellulose Syntheses

; FILE REFERENCE: B01170 US CIP

; CURRENT APPLICATION NUMBER: US/09/900,237

; CURRENT FILING DATE: 2001-07-06

; PRIOR APPLICATION NUMBER: 60/092,844

; PRIOR FILING DATE: 1998-07-14

; PRIOR APPLICATION NUMBER: PCT/US99/15871

; PRIOR FILING DATE: 1999-07-13

; PRIOR APPLICATION NUMBER: 09/720383

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 1

; LENGTH: 1221

; TYPE: DNA

; ORGANISM: Hordeum vulgare

US-09-900-237-1

```
Query Match 14.5%; Score 524.8; DB 10; Length 1221;
Best Local Similarity 75.5%; Pred. No. 7.4e-143;
Matches 652; Conservative 0; Mismatches 212; Indels 0; Gaps 0;
```

```
Qy 2394 GAAGATATTTCTACTGATTCAAGATCAGCGAAGAGGCTGGCGTTCAGTCTATTGCATG 2453
Db 5 GAGATATTTCTACTGGGTTAAATGCAAGAGGTTGGATATCAATCTACTGCATG 64
Qy 2454 CCCAAGCGCCAGCTTTTCAAGGATCTGCCCATCAATTTTCAGATCGCTGGAACCAA 2513
Db 65 CCACCAGACCTTTGTTCAAGGGTTCTGCGCCAATCAATCTCTGACCGTCTCAATCAA 124
Qy 2514 GTGCTGGGTGGGCTCTCGGTCTCTGTTGAAATCTTTTCAGCGGGCATGGCCCTTATGG 2573
Db 125 GTTCTCGGTGGGCTCTTGGGTGAGTTGAAATCTGTGTTAGCAGACATTTGCTCTATCTG 184
Qy 2574 TATGGCTACGAGGGCGCTCAAGTTCCTGGAGAGATTCGCTTACATCAACACCATT 2633
Db 185 TACAATTACGGTGGCGGGTTGAAACTTCTGGAGAGATGGCTTACATCAACACCATTGTT 244
Qy 2634 TACCACCTAACTCTCTCCGCTTTCTAGTCTATTGTATATTGGCTGTCTGTCTGCTC 2693
Db 245 TATCCAAATACATCCCTTCCACTTATGCGCTATTGTGCTTCCGTCTCTCTCTCCCTC 304
Qy 2694 ACTGGAAGTTTCATCATGCCAGAGATTAGCAACTTGGCCAGTATCTGGTTTCATTCGGCTC 2753
```

```
Db 305 ACCAACAAATTTATCATTTCCGAGATCAGTAACATATGCTGGGATGTTTATTTCTTATG 364
Qy 2754 TTCTTTTCAATTTTCGCACTGGTATCTTTGAGATAGAGTGGAGTGGTGTGGCATTTGAC 2813
Db 365 TTTGCCCTCCATCTTTGCCACGGGTATATTGGAGCTGGATGGATGGTGTGCGCATCGAG 424
Qy 2814 GAGTGGTGGAGGAATGAACAGTTCTTGGGTCAATTTGGAGGTATCTCTGCAATCTGTTGGC 2873
Db 425 GACTGGTGGAGAAACGAGCAGTTCTGGGTATTGGTGGCACATCTGCCCATCTTTTCGCA 484
Qy 2874 GTCTTTCAAGGTCTCTCTGAAGTGTCTTGAAGTGTCTTGGGTATGCTGTCACCTGTCACCTCA 2933
Db 485 GTGTTCCAGGTCTGCTGAAGGTGTGGCCGGGATTTGACACCAACTTCACGGTTACTCTCG 544
Qy 2934 AAGCTAATGACGAAGAAGCGACTTTTGTGAGCTCTACATGTTCATGTTCAGTGGAGCGCTT 2993
Db 545 AAGCAAAACGACGAGGATGGCGATTTTGTGAGTTATACGTGTCAAGTGGACCACTCTC 604
Qy 2994 CTCACTCCCTCCGACGACCATTTTGTATCATTAACATGTTGGTGTGCTGGCACCTCC 3053
Db 605 CTCACTCCCTCCGACCACTCCCTTGTGATTAACCTGGTGGGATGGTGGCAGGCATATCA 664
Qy 3054 TACGCATCAACAGTGGTTACCAATCATGGGGCCGCTCTTTGGGAAGCTCTTTCTTTGCC 3113
Db 665 TATGCCATCAACACGGGTACCAGTCTTGGGTCCACTCTTCGGAAGAGCTCTTCTTCTCA 724
Qy 3114 TTCTGGGTGATGTTGTCACTTATACCAATTCCTCAAGGGTCTTATGGCAGGCAAAACCGC 3173
Db 725 ATCTGGGTGATCCTCCATCTCTACCCCTTCTCAAGGGTCTCATGGGGAAGCAGAACCGC 784
Qy 3174 ACACCGAGGATGTCATGCTCTGGGTGTCTCTCTCGCTTCTATCTTCTCTCTCTCTCTCT 3233
Db 785 ACGCAACCAATCGTCATTTGTTGTCATCTCTCTAGCCCTCCATCTTCTCCCTCTCTCT 844
Qy 3234 GTTCGTGTTGATTCATTCATACC 3257
Db 845 GTGAAGATCGACCCCTTTCATATCC 868
```

Search completed: February 25, 2003, 22:11:45

Job time : 212 secs







PT cellulose in the pericarp, hardening the kernel for improved handling

PS ability -

XX Claim 3; Page 85-87; 88pp; English.

XX The invention relates to isolated nucleic acids encoding two cellulose

CC synthase proteins from corn. Cdp9s45 (cesa-3) and Cqrae19 (cesa-9).

CC Also disclosed are a recombinant expression cassette comprising the

CC polynucleotide (operably linked to a promoter) a host cell comprising

CC the recombinant expression cassette and a transgenic plant comprising

CC the recombinant expression cassette. The nucleic acid is useful in

CC the improvement of stalk quality for improved stand or silage, and in

CC the increased concentration of cellulose in the pericarp, hardening

CC the kernel, and thus improving its handling ability. The nucleic acids

CC may also be used as probes or amplification primers in the detection,

CC quantification or isolation of gene transcripts, as probes in detecting

CC deficiencies in the level of mRNA, for detecting gene mutations or

CC allelic variants, for monitoring up regulation of expression or changes

CC in enzyme activity in screening assays, for site directed mutagenesis,

CC and in sense or antisense suppression of one or more genes in a host

CC cell, tissue or plant. The polypeptides may be used in assays for

CC enzyme agonists or antagonists, and as immunogen or antigen to

CC obtain antibodies specifically immunoreactive with the protein.

CC The present sequence represents a corn cellulose synthase of the

CC invention.

XX Sequence 1079 AA;

SQ

Query Match 93.9%; Score 5423.5; DB 23; Length 1079;

Best Local Similarity 92.2%; Pred. No. 0;

Matches 996; Conservative 46; Mismatches 37; Indels 1; Gaps 1;

QY 1 MDGDADALSKGRHAGDVCQICADGLGTTLDGDVFTACDVCFFPCVPCYHERKEGTQA 60  
DB 1 MEGDADGVKSGRRGGQVCQICGCVGTAEAGDFTACDVCFFPCVPCYHERKDGTA 60  
QY 61 CLOCTKYKRRHGPSAIREEGEDDDTADDGSDFNYPASGTEQKOKIADRMRSWRMNTGG 120  
DB 61 CPQCKNKYKRRHGPSAIREEGEDDDTADDGSDFNYPASGTEQKOKIADRMRSWRMNTGG 120  
QY 121 SGNVGHKPYDSGEIGLSYDSGEIPRGVPSVTNSQMSGEIPGASPDHMHMSPCTNIGRR 180  
DB 121 SDGVRPYDSGEIGLTKYDSGEIPRGVPSVTNSQISGEIPGASPDHMHMSPCTNIGRR 180  
QY 181 APFPYVNHSPNPSRFFSGSIGNVANKERVQGMKQDKGAIPTMTNGTSTIAPSGRAAYDI 240  
DB 181 APFPYVNHSPNPSRFFSGSVGNVANKERVQGMKQDKGAIPTMTNGTSTIAPSGRGVGI 240  
QY 241 DASTYNNMEDALLNDETQPLSRKVPPIASSKINPYRMVIVLRLVLSIFLHYRLTNPVNR 300  
DB 241 DASTYNNMEDALLNDETQPLSRKVPPLPSSRINPYRMVIVLRLVLSIFLHYRLTNPVNR 300  
QY 301 AYPWLLSVICEIFWALSILWLDQFPKFPPIINRETYLDRLALRYDREGPSQLAAVDIFVS 360  
DB 301 AYPWLLSVICEIFWALSILWLDQFPKFPPIINRETYLDRLALRYDREGPSQLAAVDIFVS 360  
QY 361 TVDPLKEPPIVANTVLSILAVDYPVQVSVYSDGASMLTFDALAETSEFARKWVPV 420  
DB 361 TVDPMKEPPLVANTVLSILAVDYPVQVSVYSDGASMLTFDALAETSEFARKWVPV 420  
QY 421 KYDIEPRAPEFYFCOKIDYLDKQVPSFVKDRAMKREYEEFKIRINALYSKALKVPEE 480  
DB 421 KYNIEPRAPEWYFSQIDYLDKQVPSFVKDRAMKREYEEFKIRVNLGLVAKQVPEE 480  
QY 481 GHIMQDGTWPGNNTDRHPGMIQVFLHSGGLDTGEGNLPRLVYVSREKRGFQHHKAG 540  
DB 481 GHIMQDGTWPGNNTDRHPGMIQVFLHSGGLDTGEGNLPRLVYVSREKRGFQHHKAG 540  
QY 541 AMNALVRYSVAVLTNGQYMLNDCDHYINNSKAVREAMCFMDPNLGPVQVYQFPQRFDG 600  
DB 541 AMNALVRYSVAVLTNGQYMLNDCDHYINNSKALREAMCFMDPNLGRSVQVYQFPQRFDG 600  
QY 601 IDNRDRIANRNTVFEDINLRGLDQGPVYVGTGCFVNRNTAIYGEPPKAKKPGFLASL 660

DB 601 IDNRDRIANRNTVFEDINLRGLDQGPVYVGTGCFVNRNTAIYGEPPKAKKPGFLASL 660  
QY 661 CGGKKASKSKSRSSDKKSKNR 720  
DB 661 CGGKKASKSKK -GSDKKSKSR 719  
QY 721 RFGSAAFVASTLMEYGGVPOS 780  
DB 720 RFGSAAFVASTLMEYGGVPOS 779  
QY 781 ILTFGKHARGWRSVYCMKPR 840  
DB 780 ILTFGKHARGWRSVYCMKPR 839  
QY 841 YGGRKFLERPAYINTTIYPL 900  
DB 840 YGGRKFLERPAYINTTIYPL 899  
QY 901 SIFATGILEMWSGVGIDEMW 960  
DB 900 SIFATGILEMWSGVGIDEMW 959  
QY 961 NDEEGDFAELYMFKWTLLIP 1020  
DB 960 SDEGDFAELYMFKWTLLIP 1019  
QY 1021 VIVHLYPFLKGLMGRQNRTP 1080  
DB 1020 VIVHLYPFLKGLMGRQNRTP 1079

RESULT 2

AAY84108

ID AAY84108 standard; Protein 7 AA.

XX AAY84108;

XX 03-JUL-2000 (first entry)

XX Amino acid sequence of a maize cellulose synthase.

XX Maize; cellulose synthase; cell quality; stand; silage; cellulose;

XX transgenic plant; plant breeding marker.

XX Zea mays.

XX Key Location/Quarters

XX Misc-difference 494 /note= "entry by MGG"

XX WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US1876

XX 17-AUG-1998; 98US-009682

XX (PION-) PIONEER HI-BRED IN

XX Dhugga KS, Helentjaris TC

XX WPI; 2000-224343/19.

XX N-PSDB; AA299494.

XX New genes which encode maize cellulose synthase polypeptides in plants

XX useful for modulating the expression of cellulose synthase in plants

XX and to produce transgenic plants expressing the novel protein

XX Claim 15; Page 95-97; 119p

XX The present sequence represents a maize cellulose synthase polypeptide.

The cellulose synthase can be used for the improvement of stalk quality for improved stand or silage. It also provides an increased concentration of cellulose in the pericarp, hardening the kernel and improving its handling ability. The sequences are used to produce transgenic plants and seeds expressing the cellulose synthase. The polynucleotide is used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Query Match 93.7%; Score 5415.5; DB 21; Length 1077;  
Best Local Similarity 92.3%; Pred. NO. 0;  
Matches 997: Conservative 46; Mismatches 34; Indels 3; Gaps 2;

QY	1	MDGDADALASGRHAGADVCOICADGLGTLTLDGDVFTACDVCRRPVCPCYEHREKGTQA	60
Db	1	MESGDAGVKSRRGGGQVQICGDGVGTTAEGDVFACDVCGPVCRPCYETERDKDTQA	60
QY	61	CLOCKTKYKRRHSIPAIRBEGDDTDADDSGFNYFASGTEDQKQIADRMESWRMTGG	120
Db	61	CPQCKTKYKRRHKSIPAIRBEGDDTDAD--SDFNYLASNEQKQIADRMESWRNVGG	118
QY	121	SGNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVTNSOMSGEIPCASPDHMHMSPTGNISRR	180
Db	119	SGDVGRPKYDSGEIGLTKYDSGEIPRGYTPSVTNSQISGEIPCASPDHMHMSPTGNIGKR	178
QY	181	APPPYVNHSPNPSREPSGSGTGNVAKWERVDGWMKMDKGAIPMTNGTSTAPSEGRAATDI	240
Db	179	APPPYVNHSPNPSREPSGSGTGNVAKWERVDGWMKMDKDGTIPTNGTSTAPSEGRGVGDI	238
QY	241	DASTEYNMEDALLNDETROPLSRKVPISASKINPYRMVTLRLVLVLSIFELHYRLTNPVRN	300
Db	239	DASTDYNMEDALLNDETROPLSRKVPLPSSRINPYRMVTLRLVLVLSIFELHYRIITPNRN	298
QY	301	AYPLWLLSVICETWFALSWILDOPFKWFPINRETYLDRLALRYDRBEGEPSQAAVDIFVS	360
Db	299	AYPLWLLSVICEIWFALSWILDOPFKWFPINRETYLDRLALRYDRBEGEPSQAAVDIFVS	358
QY	361	TVDPLEPPITVTANTVLSILAVDYPVDKYSVCYVDDGASMLFADALAEHSEARKWVPV	420
Db	359	TVDPMEPPITVTANTVLSILAVDYPVDKYSVCYVDDGAAMLFTDALAEHSEARKWVPV	418
QY	421	KKYDIEPRAPEFFCOKIDYLDKQVQSFVKDRAMKREYEEFKIRINALVSALKALVPEE	480
Db	419	KKYNIEPRAPEWFSQIDYLDKQVHPSFVKDRAMKREYEEFKVRVNGLVAKAQKVPPE	478
QY	481	GWIMQDGTWPGNNTRDHPGMIQVFIHGSGGLDTEGNEUPLRVLYVVSREKRPQFQHHKAG	540
Db	479	GWIMQDGTWPGNNTDHPGMIQVFIHGSGGLDTEGNEUPLRVLYVVSREKRPQFQHHKAG	538
QY	541	AMNALVRVSAVLNTNGVYMLNLDGDHYINNSKAVREAMCFLMDPNLGPOVCYVQVFPORFDG	600
Db	539	AMNALVRVSAVLNTNGVYMLNLDGDHYINNSKALREAMCFLMDPNLGRSVYVQVFPORFDG	598
QY	601	IDRNDRYANRNTVFFDINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFASL	660
Db	599	IDRNDRYANRNTVFFDINLRGLDGIQGVYVGTGCVFNRTALYGYEPPKQKKGFLSSL	658
QY	661	CGGKKKASKKRSKSSKKSNKHVDSSVPFNLEDIEEGVEGAGFODEKSVLMSQMSLEK	720

used for modulating, preferably increasing, the level of the synthase in a plant cell. The plants are preferably monocots. The polynucleotide is also used as a probe or primer in the detection quantitation or isolation of gene transcripts. The probes are useful in detecting deficiencies in the level of mRNA in screenings for desired transgenic plant, for detecting mutations in the gene, for monitoring upregulation of expression or changes in enzyme activity in screening assays of compounds, for detection of any number of allelic variants of the gene, or for use as molecular markers in plant breeding programs. The isolated nucleic acids of the present invention can also be used for recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.

Query Match 93.7%; Score 5415.5; DB 21; Length 1077;  
Best Local Similarity 92.3%; Pred. No. 0;  
Matches 997; Conservative 46; Mismatches 34; Indels 3; Gaps 2;

QY 1 MGDADALKSGRHGAGDVCQICADGLTGLDGVTAACVCRPCRCYEHKEGTA 60  
DB 1 MEGDADGVKSGRRGGQVCQICGDDGVTGTAEGDFAACDVCVCRPCYERKDTQA 60  
QY 61 CLOCKTKYKRRHSGPAIRGEEDDDADDGSDFNYPASGTEDOKKIADRMESWRNTGG 120  
DB 61 CPOCKTKYKRRHSGPAIRGEEDDDADD--SDFNILASGNEQKOKIADRMESWRNVGG 118  
QY 121 SGNVGHKYDSGEIGLSKYDSGEIPRGYVPSVTNSQMSGEIPGASPDHMSPTGNISRR 180  
DB 119 SGNVGRPKYDSGEIGLTKYDSGEIPRGYVPSVTNSQISGEIPGASPDHMSPTGNIGKR 178  
QY 181 APFPYVNSPNRPFSSGIGNVANKERVGVKQDKGAIPTWNGTSTAPSEGRAADTI 240  
DB 179 APFPYVNSPNRPFSSGIGNVANKERVGVKQDKGAIPTWNGTSTAPSEGRGVGDI 238  
QY 241 DASTYNNMEDALLNDETROPLSRKYPVIASSKINPYRMVILRLVLSIFLHYRLNPVRN 300  
DB 239 DASTYNNMEDALLNDETROPLSRKYPVLPSSRINPYRMVILRLVLSIFLHYRIINPVN 298  
QY 301 AYPLMLLSVICBIWFALSWILDQFPKWPINRETYDLRLALRYDREGESQAAVDIFVS 360  
DB 299 AYPLMLLSVICBIWFALSWILDQFPKWPINRETYDLRLALRYDREGESQAAVDIFVS 358  
QY 361 TVDPLKEPIVTANTVLSLAVDYPVVKVSCVSDGASMLTFDALAETSEFARKWVPV 420  
DB 359 TVDPMKEPLVTANTVLSLAVDYPVVKVSCVSDGAAMLTFDALAETSEFARKWVPV 418  
QY 421 KKYDIEPRAPEFYFCQKIDYLDKQVQSFVKDRRAMKREYEEFKIRINALVSKALKVPEE 480  
DB 419 KKYDIEPRAPEFYFSQKIDYLDKQVHPSFVKDRRAMKREYEEFKVRVNGLVAKQVPEE 478  
QY 481 GHIMQDGPWPNNTRDHPGTMTOVFLHSGGLDTGNEPLRLVYVSREKRPQFHHKAG 540  
DB 479 GHIMQDGPWPNNTRDHPGTMTOVFLHSGGLDTGNEPLRLVYVSREKRPQFHHKAG 538  
QY 541 ANNALVRYSAVLTNGOYMLNLDCHYINNSKAVRAMCFMIDPDLGPOVCYVQFQRPDQ 600  
DB 539 ANNALVRYSAVLTNGOYMLNLDCHYINNSKALRAMCFMIDPDLGSRVCYVQFQRPDQ 598  
QY 601 IDNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVNFRNTALYGEYEPPIKAKKPGFLASL 660  
DB 599 IDNDRYANRNTVFDDINLRGLDGIQGPVYVGTGVNFRNTALYGEYEPPIKAKKPGFLSSL 658  
QY 661 CGGKKKASKRSKRSKSKKSNKHVDSSVPVFNLEDIEEGVEGAGDDEKSVLMSQMSLEK 720  
DB 659 CGGRKKASKSKK-GSDKKKSKQKVDSSVPVFNLEDIEEGVEGAGDDEKSLMSQMSLEK 717  
QY 721 RFGQNAFVASTLMYGGVQPSSTPESLLKEAHIVISCGYEDKSWGTGIGWYGSVTEB 780  
DB 718 RFGQNAFVASTLMYGGVQPSATPESLLKEAHIVISCGYEDKTEWGTGIGWYGSVTEB 777

QY 781 ILTGFKMHARGWRSVYVCMKRP  
DB 778 ILTGFKMHARGWRSVYVCMKRP  
QY 841 YGRLKFLERFAYINTTIYPLT  
DB 838 YGRLKFLERFAYINTTIYPLT  
QY 901 SIFATGILEMRSGVGIDEMWR  
DB 898 SIFATGILEMRSGVGIDEMWR  
QY 961 NDEEGFAELYMFKWTTLLIPP  
DB 958 SDEGDFAELYMFKWTTLLIPP  
QY 1021 VIVHLYPFLKGLMGRNRTPTI  
DB 1018 VIVHLYPFLKGLMGRNRTPTI

RESULT 4  
AA84119  
ID AAY84119 standard; Protein;  
XX AC AAY84119;  
XX DT 03-JUL-2000 (first entry)  
XX DE Amino acid sequence of a ma  
XX KW Maize; cellulose synthase;  
XX KW transgenic plant; plant bre  
XX OS Zea mays.  
XX FH Key  
XX FT Misc-difference 494  
XX XX Location/Qu  
XX PN WO200009706-A2.  
XX PD 24-FEB-2000.  
XX PF 16-AUG-1999; 99WO-US18760  
XX PR 17-AUG-1998; 98US-0096822  
XX PA (PION-) PIONEER HI-BRED INT  
XX PI Dhugga KS, Helentjaris TG,  
XX DR WPI: 2000-224343/19.  
XX DR N-PSDB; AA299527.  
XX PT New genes which encode maiz  
XX PT useful for modulating the ex  
XX PT and to produce transgenic p  
XX PS Claim 15; Page 181-183; 119  
XX XX

The present sequence repres  
The cellulose synthase can  
for improved stand or silag  
of cellulose in the pericar  
handling ability. The sequ  
and seeds expressing the ce  
used for modulating, prefer  
in a plant cell. The plants  
is also used as a probe or  
isolation of gene transcrip  
deficiencies in the level o  
a maize cellulose synthase polypeptide.  
ed for the improvement of stalk quality  
also provides an increased concentration  
rdening the kernel and improving its  
are used to produce transgenic plants  
se synthase. The polynucleotide is  
increasing, the level of the synthase  
preferably monocots. The polynucleotide  
r in the detection quantitation or  
the probes are useful in detecting  
A in screenings for desired transgenic



CC recombinant expression of their encoded polypeptides or for use as immunogens in the preparation and/or screening of antibodies. The proteins can be employed in assays for enzyme agonists or antagonists of enzyme function or for use of immunogens or antigens to obtain antibodies specifically immunoreactive with a protein.									
XX	Sequence	1076	AA;						
QY	Query Match	89.0%;	Score 5143;	DB 21;	Length 1076;				
Db	Best Local Similarity	87.6%;	Pred. No. 0;						
Matches	947;	Conservative	69;	Mismatches	59;	Indels	6;	Gaps	6;
QY	1	MDGDALKSGRHGAGDVCQICADGLTLDGVDFTACDVCRCPCYEHKEGTQA	60						
Db	1	MDG-GDATSGKHVAGQVQICGDGVTAGDGLFTACDVCRCPCYEHKEGTQA	59						
QY	61	CLQCKTKYKRRHSGPAIRGEGDDTDADDGSDPNYPASGTEDQKQIADRMRSRMNTGG	120						
Db	60	CPQCKTKYKRRHSGPPVHGEENEDVDADDSDYNYQSGNQDQKQIAERMLTWRNTRSG	119						
QY	121	SGNVGHPKYDSGEIGLSKYSDBEIPRGYPVSVTNSQSGEIPGASPDHMHMSTGNTISRR	180						
Db	120	S-DIGLAKYDSGEIGHKYSDBEIPRGYIPSLTHSQISGEIPGASPD-HMSPVGNIGRR	177						
QY	181	A-PFVYNHSPNPSRFSGSIGNVANKERVWGKMKQDKGAIPMTNGTSTAPSEGAARD	239						
Db	178	GHQFPYVNHSPNPSRFSGSLGNVANKERVWGKMK-DKGAIPMTNGTSTAPSEGRVAD	236						
QY	240	IDASTYNNMEDALLNDTQPLSRKVPYIASSKINPYRMVILVRLVLSIFLHVRLNPNVR	299						
Db	237	IDASTYNNMEDALLNDTQPLSRKVPYIPSSRNPNRMVILVRLVLCIFLRYRITHPNV	296						
QY	300	NAYPLWLLSVICEIWFALSWMILDOFPKWPFINRETYLDRLALRYDREGPSQLAAYDIFV	359						
Db	297	NAYPLWLLSVICEIWFALSWMILDOFPKWPFINRETYLDRLALRYDREGPSQLAPVDIFV	356						
QY	360	STVDPLKEPPIYNTANTVLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPF	419						
Db	357	STVDPMKEPPIYNTANTVLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPF	416						
QY	420	VKKYDIEPRAPEFYFCQKIDYLDKQVPSFVKDRRAKREYEFKIRINALVSKALKVPE	479						
Db	417	CKKYNIEPXAPEWYFAQIDYLDKQVTSFVKERRAKREYEFKIRINCLVAKAKVPE	476						
QY	480	EGWIMQDGPWGNNTRDHPGMIVQVFLHSGGLDTEGNEPLRVLYVVSREKRPQFQHHKA	539						
Db	477	EGWIMQDGPWGNNTRDHPGMIVQVFLHSGGLDTEGNEPLRVLYVVSREKRPQFQHHKA	536						
QY	540	GAMNALVRYSAVLITNGQYMLNLDCHYINNSKAVREAMCFMDPNLGPQVCYVQFPQRED	599						
Db	537	GAMNALVRYSAVLITNGQYMLNLDCHYINNSKALREAMCFMDPNLGRNVYVQFPQRED	596						
QY	600	GIDRNDRYANRNTVFDDINLRGLDGTGGPVYVGTGCVFNRTAIFYEPPITAKKPKGFLAS	659						
Db	597	GIDRNDRYANRNTVFDDINLRGLDGTGGPVYVGTGCVFNRTAIFYEPPVKKKPKGFFSS	656						
QY	660	LCGGKKKASKKRSDSKKSNKHVDSSVPVFNLEDIEEGVAGAGFDDDEKSLYMSQMSLE	719						
Db	657	LCGGRRKTKSKSK-SSEKKKSHRHADSSVPVFNLEDIEEGSQFDDDEKSLYMSQMSLE	715						
QY	720	KREGQSAAFVASTLMIEYGGVQSSPESLLKEAIIHVISGVEDKSEWGTGIEGWIYSVTE	779						
Db	716	KRFQGSVFVASTLMIEYGGVQSSPESLLKEAIIHVISGVEDKTDWGTGIEGWIYSVTE	775						
QY	780	DILITGFKMHARGWRSYCYMKPAPKFGSAPINLSDRNLQVLRWALGSVELFSRHCPWY	839						
Db	776	DILITGFKMHARGWRSYCYMKPAPKFGSAPINLSDRNLQVLRWALGSIELFSRHCPWY	835						
QY	840	GYGGRUKFLERFAYINTTITPLTSLPLLYCILPAICLLTGKTFIMPEISNLASIFIALF	899						
Db	836	GYGGRUKFLERFAYINTTITPLTSLPLLYCILPAICLLTGKTFIPKISNLESVWFISLF	895						
QY	900	LSIFATGILEMRWSGVGIDENWRNEQFWTIGGISAHLPFAVFOGLLKVLAGIDNTFTVTSK	959						

Db	896	ISIFATGILEMRWSGVGIDENWRNEQFWTIGGISAHLPFAVFOGLLKVLAGIDNTFTVTSK	955						
QY	960	ANDEEGDFAELYMKFTTLLIP							
Db	956	ATDEEGDFAELYMKFTTLLIP							
QY	1020	WVIVHLYPFLKGLMGRQNRTP							
Db	1016	WVIVHLYPFLKGLMGRQNRTP							
QY	1080	C 1080							
Db	1076	C 1076							
RESULT	6								
AAG48734									
ID	AAG48734	standard; Protein:	AA.						
XX	AAG48734;								
XX	18-OCT-2000	(first entry)							
XX	Arabidopsis thaliana	protein	segment SEQ ID NO: 61574.						
XX	Protein identification; sig	transduction pathway; metabolic pathway;							
XX	hybridisation assay; genet	cloning; gene expression control; promoter;							
XX	termination sequence.								
XX	Arabidopsis thaliana.								
XX	EP1033405-A2.								
XX	06-SEP-2000.								
XX	25-FEB-2000;	2000EP-0301439							
XX	25-FEB-1999;	99US-0121825							
XX	05-MAR-1999;	99US-0123180							
XX	09-MAR-1999;	99US-0123548							
XX	23-MAR-1999;	99US-0125788							
XX	25-MAR-1999;	99US-0126264							
XX	29-MAR-1999;	99US-0126785							
XX	01-APR-1999;	99US-0127462							
XX	06-APR-1999;	99US-0128234							
XX	08-APR-1999;	99US-0128714							
XX	16-APR-1999;	99US-0129845							
XX	19-APR-1999;	99US-0130077							
XX	21-APR-1999;	99US-0130449							
XX	23-APR-1999;	99US-0130510							
XX	28-APR-1999;	99US-0130891							
XX	30-APR-1999;	99US-0131449							
XX	30-APR-1999;	99US-0132048							
XX	04-MAY-1999;	99US-0132407							
XX	05-MAY-1999;	99US-0132484							
XX	06-MAY-1999;	99US-0132485							
XX	06-MAY-1999;	99US-0132486							
XX	07-MAY-1999;	99US-0132487							
XX	07-MAY-1999;	99US-0132863							
XX	11-MAY-1999;	99US-0134256							
XX	14-MAY-1999;	99US-0134218							
XX	14-MAY-1999;	99US-0134219							
XX	14-MAY-1999;	99US-0134221							
XX	14-MAY-1999;	99US-0134370							
XX	18-MAY-1999;	99US-0134768							
XX	19-MAY-1999;	99US-0134941							
XX	20-MAY-1999;	99US-0135124							
XX	21-MAY-1999;	99US-0135353							
XX	24-MAY-1999;	99US-0135629							
XX	25-MAY-1999;	99US-0136021							
XX	27-MAY-1999;	99US-0136392							
XX	28-MAY-1999;	99US-0136782							

PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139482.  
PR 18-JUN-1999; 99US-0139483.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149358.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 23-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.



Query Match		78.4%;	Score 4530;	DB 21;	Length 1054;
Best Local Similarity		78.6%;	Pred. No. 0;		
Matches 839;		Conservative 96;	Mismatches 104;	Indels 26;	Gaps 8;
Qy	19	COICADGLTGTACDVCPCPCYEHREKGTGACLOCKTKYKRHRGSPAIR	78		
Db	9	COICSDNVGKTVGDGFVACDPCPCYEHREKGTGACLOCKTKYKRHRGSPAIR	68		
Qy	79	GEEDDDTADGGS-DFNYPASGTEDQKOKIADRMRSRMNTGGSGNVGHPKYDSEIGLS	137		
Db	69	GDKDEDGLADSGTVEFNYP-----QKEKISERMLGHLTRGKGEEMGEPOYDK-----	116		
Qy	138	KYDSSEIPRGVVPSTNSQ-MSGEIPGASPDHMHMSPGNSRRAPPY-VNHSPNSPRE	195		
Db	117	-----EVSHNLPLRLTSQDTSGGFSAAPELRSVSTAGKRLPYSSDYNQSPNRIV	171		
Qy	196	FGSIGNVAMKRDVGMKMDKGAIPMTNGTSIAPSEGRAATDIDASTETYNMEDALLND	255		
Db	172	DPVGLGNVAKRVDGWMKMKQKNGTGPV--STQAASE-RGVVDIDASTDILADEALLND	227		
Qy	256	ETRQPLSRKVIASKINPYRMVIVLRLVLSIFLHYRLTNPNVFNAYPLWLLSVICEIWF	315		
Db	228	EARQPLSRKVIASKINPYRMVIVLRLVLSIFLHYRLTNPNVFNAYPLWLLSVICEIWF	287		
Qy	316	ALSWILDQFPKWFPLNRETYDRLALRYDREGPEPSLAAVDIFVSTVDPLKEPPIVTANT	375		
Db	288	ALSWILDQFPKWFPLNRETYDRLALRYDREGPEPSLAAVDIFVSTVDPLKEPPIVTANT	347		
Qy	376	VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPVKYDIEPAPPEYFC	435		
Db	348	VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWVPVKYDIEPAPPEYFC	407		
Qy	436	OKIDYLDKDVOPSFVKDRAMKREVEPKIRINALVSKALVPEEGTIMODGTWPNGNNT	495		
Db	408	AKIDYLDKDVOPSFVKDRAMKREVEPKIRINALVSKALVPEEGTIMODGTWPNGNNT	467		
Qy	496	RHDPGMIQVFLGSHGGLDTEGNELRLVYVSREKRPQFQHKKAGAMNALVRVSAVLTN	555		
Db	468	RHDPGMIQVFLGSHGGLDTEGNELRLVYVSREKRPQFQHKKAGAMNALVRVSAVLTN	527		
Qy	536	QYMLNLCDDHYINNSKALREAWCFMPLDNLGPQVCYVQFQFQFQFQFQFQFQFQFQF	615		
Db	528	PFILNLCDDHYINNSKALREAWCFMPLDNLGPQVCYVQFQFQFQFQFQFQFQFQFQF	587		
Qy	616	DINRLGLDGIQGVYVGTGCVFNRTAIYGYEPPPIKA--KKPGFLASLCGGKKKSKSKR	673		
Db	588	DINRLGLDGIQGVYVGTGCVFNRTAIYGYEPPPIKA--KKPGFLASLCGGKKKSKSKR	647		
Qy	674	SSDKKSKNKHVDSSVPVFNLEDEIEGVGAGFDDDEKSVLMSQMSLEKRFQGSAAFAVSTL	733		
Db	648	ESDKKSKNKHVDSSVPVFNLEDEIEGVGAGFDDDEKSVLMSQMSLEKRFQGSAAFAVSTL	707		
Qy	734	MEYGVPOSSPTSELKKAHVISCGYEDKGEWGTGIGWIGSVTDTLTGFKMHARGWR	793		
Db	708	MENGCVPSATPENLLKKAHVISCGYEDKSDWGEIGWIGSVTDTLTGFKMHARGWR	767		
Qy	794	SYTCMPKRPATKGSAPINLSRLNQVLKALGVSVEILFSRHCPLWYGYGRLKFLERFAY	853		
Db	768	SYTCMPKLPATKGSAPINLSRLNQVLKALGVSVEILFSRHCPLWYGYGRLKFLERFAY	827		
Qy	854	INTTITPLTSLPLLYVCILPAICLLTGKFMPEISNLSIASIFIALFLSIFATGILEMWS	913		
Db	828	VNTTIYPTISPLMICYFLPAVFTNQFLIPQISNLSIASIFIALFLSIFATGILEMWS	887		
Qy	914	GVGIDEMWRNQFWIGGISAHFLFAVFGGLKVLGAGIDTNTVTTSKANDDEGDAEALYMF	973		
Db	888	GVGIDEMWRNQFWIGGISAHFLFAVFGGLKVLGAGIDTNTVTTSKANDDEGDAEALYMF	947		
Qy	974	KWTTLLIPPTTILLINMVGVAAGTSYAINSGYQSGWPLFGKLFPAFWIVVHLYPFLKGLM	1033		
Db	948	KWTTLLIPPTTILLINMVGVAAGTSYAINSGYQSGWPLFGKLFPAFWIVVHLYPFLKGLM	1007		
Qy	1034	GRNRTPTTIVIVWAVLLASIFSLWVRVDPPTTRLAGPNIQTCGINC	1080		

Db	1008	GRGNRTPTTIVVWSVLLASIFS	1054	: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
----	------	------------------------	------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------



```
QY 138 KYDSGEIPRGVYPSVTNSQ-MSGEIPGASPDHMHMSPTGNISRRAPPY-VNHSPNPSRE 195
Db 128 -----EVSHNHLPLRTSRQDTSGESAASPERLSVSTIAGGKRLPYSSDVNQSPNRRIV 182
QY 196 FSGSTGNVANKRVDGKWKQDKGAIPMTNGTSHAPSEGRATDIDASTEYWNEDALLND 255
Db 183 DPGVLGNVANKRVDGKWKQDKGVPV---STQASE-RGGVDIDASTDILLADEALLND 238
QY 256 ETRQPLSRKVPATSSKINPYRMVIVLRLVLSIFLHYRLTNPNVFNAYPLWLLSVICIEWF 315
Db 239 EARQPLSRKVSIPSSKINPYRMVILRLVLCFLHYRLTNPNVFNAYPLWLLSVICIEWF 298
QY 316 ALSWILDQFPKWFPIINRETYLDRALRYDREGEPSQAAVDIFVSTVDPLKEPPIVANT 375
Db 299 ALSWILDQFPKWFPIINRETYLDRALRYDREGEPSQAAVDIFVSTVDPLKEPPIVANT 358
QY 376 VLSILAVDYPVKVSCYVSDGASMLTFDALAETSEFARKWVPFVKKYDIEPRAPEYFC 435
Db 359 VLSILAVDYPVKVSCYVSDGAAMLSFEALAESEFARKWVPFVKKYDIEPRAPEYFA 418
QY 436 QKIDYLDKQVPSFVKDRAMKREVEEFKIRINALVSKALKVPEEGIMQDGTTPWGNNT 495
Db 419 AKIDYLDKQVPSFVKDRAMKREVEEFKIRINALVSKALKVPEEGIMQDGTTPWGNNT 478
QY 496 RDHPGMIQVFLGHSGGLDTEGNELPLRVYVSRKPRGFQHHKKAGAMNALVRVSAVLTNG 555
Db 479 RDHPGMIQVFLGHSGGLDTEGNELPLRVYVSRKPRGFQHHKKAGAMNALVRVSAVLTNG 538
QY 556 QYMLNLDCHYTNNSKAVPEACFLMDPNLGPQVCYVQPPQDFDGDIDRNDRYANRNTVFF 615
Db 539 PFILNLDCHYTNNSKALREACFLMDPNLGPQVCYVQPPQDFDGDIDRNDRYANRNTVFF 598
QY 616 DNLRLGLDGIQGVYVGTGVNRYAIYGYEPPIKA--KKPGFLASLCGGKKKASKSKR 673
Db 599 DNLRLGLDGIQGVYVGTGVNRYAIYGYEPPIKVYKHKKPSLLSKLGGSRKKNSKAKK 658
QY 674 SSDKSKNHVDSVPVFNLEIDEEGVAGDEDEKSVLMQMSLEKRFQGSAAVFASTL 733
Db 659 ESDKKKSGRHTDSTVPVFNLDDEEGVEGAGDEDEKALLMSQMSLEKRFQGSAAVFASTL 718
QY 734 MEYGVQSSQTESLILKEAHIWISCYEDKSEWGTETIGWYGSVTEDILTGFKMHARGWR 793
Db 719 MENGVPPSPATENLILKEAHIWISCYEDKSDWMEIGWYGSVTEDILTGFKMHARGWR 778
QY 794 SYCYMKPFAFGSAPINLSRLNQVLRWALGSVEILFSRHCPIMWYGGRLKFLERFAY 853
Db 779 SYCYMKPFAFGSAPINLSRLNQVLRWALGSVEILFSRHCPIMWYNGRLKFLERFAY 838
QY 854 INTTYPLSLPLVYCIIPALCLLTGKFMPEISNLASITWFTALFSLFATGILEMRWS 913
Db 839 VNTTYPIITSIPLLMYCTLPVAVCLFTNQFIIPQISNIASITWFTALFSLFATGILEMRWS 898
QY 914 GVGIDSWRNEQFWTGGISAHFAVFOGLLKVLGIDNTFTVTSKANDEEGDFAELYMF 973
Db 899 GVGIDSWRNEQFWTGGISAHFAVFOGLLKVLGIDNTFTVTSKANDEEGDFAELYLF 958
QY 974 KWTLLIPPTLIIINMGVWAGTSVAINSYGQSWGPGFKLPFAFWIVHLYPFLKGLM 1033
Db 959 KWTLLIPPTLIIINMGVWAGTSVAINSYGQSWGPGFKLPFAFWIVHLYPFLKGLM 1018
QY 1034 GRONRPTIVYVAVLSTFSLWVRVDPFTRLAGNITQTCGINC 1080
Db 1019 GRONRPTIVYVAVLSTFSLWVRVDPFTRLAGNITQTCGINC 1065

RESULT 8
ID ABB93356 standard; Protein; 1065 AA.
XX
AC ABB93356;
XX
DT 31-MAY-2002 (first entry)
XX
```

```
DE Herbicidally active polypeptide
XX Herbicidal; plant; agricultural
KW Herbicidal; plant; agricultural
OS Arabidopsis thaliana.
XX WO200210210-A2.
PN 07-FEB-2002.
XX 28-AUG-2001; 2001WO-EP09892
PR 28-AUG-2001; 2001WO-EP09892
XX (FARB ) BAYER AG.
PA Tietjen K, Weidler M;
PI WPI; 2002-269010/31.
DR Identifying plant target proteins comprising
XX PT comprising aligning and comparing nucleic acid
PT from plant with nucleic acid sequences from non-plant
PT organisms -
XX Claim 5; SEQ ID NO 2567; 26
PS The invention relates to identifying target proteins
XX (ABB90790-ABB94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences having an E-value
CC suitable search parameters, greater by a factor of 3 than
CC sequences are selected. The peptides or nucleic acids encoding them
CC are useful for identifying herbicides.
XX are useful as herbicides.
SQ Sequence 1065 AA;
Query Match 78.4%;
Best Local Similarity 78.6%;
Matches 839; Conservative
QY 19 CQICADGLGTTLDGVDFTACDV 1065
Db 20 CQICSDNVGKTVDDGDFVACDI 1065
QY 79 GEGDDTDADDGS-DEFNYPASG 1065
Db 80 GKDEGLADEGIVENYPP- 1065
QY 138 KYDSGEIPRGVYPSVTNSQ-MS 1065
Db 128 -----EVSHNHLPLRTSRQDT 1065
QY 196 FSGSIGNVANKRVDGKWKQDK 1065
Db 183 DPGVLGNVANKRVDGKWKQDK 1065
QY 256 ETRQPLSRKVPATSSKINPYRM 1065
Db 239 EARQPLSRKVSIPSSKINPYRM 1065
QY 316 ALSWILDQFPKWFPIINRETYLD 1065
Db 299 ALSWILDQFPKWFPIINRETYLD 1065
QY 376 VLSILAVDYPVKVSCYVSDG 1065
Db 359 VLSILAVDYPVKVSCYVSDG 1065
QY 436 QKIDYLDKQVPSFVKDRAMKRE 1065
```

```
SEQ ID NO 2567.
herbicide.
s for herbicidally active compounds,
g nucleic acid or amino acid sequences
amino acid sequences from non-plant
organisms -
Sequence Listing: English.
ying target proteins
ally active compounds, comprising
cid or amino acid sequences from plant
sequences from non-plant organisms using
e plant sequences having an E-value
e E-value of most similar non-plant
peptides or nucleic acids encoding them
ators. The identified modulators are
are 4530; DB 23; Length 1065;
d. No. 0;
Mismatches 104; Indels 26; Gaps 8;
CRPCYEHKKEGTQACLOCKTKYKRRHSGPAIR 78
CRPCYEHKKGQSCPCQCKTKYKRLKSGPAIR 79
SOKIADMRSMRMNTGGSGNVGHPKYSGEIGLS 137
EKISERMLGWHLTRGKEEGEPQYDK----- 127
ASPDHMHMSPTGNISRRAPPY-VNHSPNPSRE 195
ASPERLSVSTIAGGKRLPYSSDVNQSPNRRIV 182
MTNGTSTAPSEGRAATDIDASTEYWNEDALLND 255
V-----STQASE-RGGVDIDASTDILLADEALLND 238
LVVLSIFLHYRLTNPNVFNAYPLWLLSVICIEWF 315
LVILCLFLHYRLTNPNVFNAYPLWLLSVICIEWF 298
HYDEGEPSQAAVDIFVSTVDPLKEPPIVANT 375
HYDEGEPSQAAVDIFVSTVDPLKEPPIVANT 358
FDALAESEFARKWVPFVKKYDIEPRAPEYFC 435
SFESLAETSEFARKWVPFVKKYDIEPRAPEYFA 418
FKIRINALVSKALKVPEEGIMQDGTTPWGNNT 495
```



Db	599	DINLRGLDGIQGVYVGTGCVNFNTALVGYEPPPIKVHKHKKPSSLLSKLGGSRKNKSKAKK	658	PR	06-MAY-1999;	99US-0134246
Qy	674	SSDKKKSNKHVDSPVFNLEIDIEEGVEGAGFDDKESVLMSQMSLEKRFQCSAAVASTL	733	PR	07-MAY-1999;	99US-0134255
Db	659	ESDKKKSGRHTDSTVPVFNLDIDIEEGVEGAGFDDKESVLLMSQMSLEKRFQCSAAVASTL	718	PR	11-MAY-1999;	99US-0134255
Qy	734	MEYGGVQSSTPSELLEAHVHISCGVEDKSEWGTGICWYGSVTEIDILGCFKMHARGWR	793	PR	14-MAY-1999;	99US-0134277
Db	719	MENGGVPPSATPFLKEAHVHISCGVEDKSDGMEIGWYGSVTEIDILGCFKMHARGWR	778	PR	14-MAY-1999;	99US-0134376
Qy	794	SVYCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPILWYGGRLKFLERFAY	853	PR	18-MAY-1999;	99US-0134944
Db	779	SIYCMKPLPAKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPILWYGGRLKFLERFAY	838	PR	19-MAY-1999;	99US-0135124
Qy	854	INTIYPLTSLPLVYCYLPAICLLTGKFMPEISNLASITWFTALFISFATGILEMRWS	913	PR	21-MAY-1999;	99US-0135355
Db	839	VNTIYPTISPLMYCTLAVCLFTNQFIQISNTIASITWFTSLFSLFISFATGILEMRWS	898	PR	24-MAY-1999;	99US-0135622
Qy	914	GVGIDEWNRNQFWVIGGISAHLFAVFGILLKVLGIDTNTFTVTSKANDDEGDAEILYMF	973	PR	25-MAY-1999;	99US-0136024
Db	899	GVGIDEWNRNQFWVIGGISAHLFAVFGILLKVLGIDTNTFTVTSKANDDEGDAEILYMF	958	PR	27-MAY-1999;	99US-0136359
Qy	974	KWTTLLIPPTTILINMVGVSAGTSYAINSGYSQSWGFLGKLFFAFWVIVHLYPFLKGLM	1033	PR	28-MAY-1999;	99US-0136788
Db	959	KWTTLLIPPTTLLVNLGVVAGVSYAINSGYSQSWGFLGKLFFAFWVIVHLYPFLKGLM	1018	PR	01-JUN-1999;	99US-0137224
Qy	1034	GRQNRTPITVILVAVILLASFLLWVRVDEFTTFLAGNPOTCGINC	1080	PR	03-JUN-1999;	99US-0137500
Db	1019	GRQNRTPITVIVVLLASIFSLWVRIDPFTTSKVTGPDILECGINC	1065	PR	04-JUN-1999;	99US-0137500
RESULT 10						
AAG48735						
ID		AAG48735 standard; Protein; 959 AA.				
AC		AC				
XX		AC				
XX		AC				
DT		18-OCT-2000 (first entry)				
XX		XX				
DE		Arabidopsis thaliana protein fragment SEQ ID NO: 61575.				
KW		Protein identification; signal transduction pathway; metabolic pathway;				
KW		hybridisation assay; genetic mapping; gene expression control; promoter;				
XX		termination sequence.				
XX		XX				
OS		Arabidopsis thaliana.				
PN		EPI033405-A2.				
PD		06-SEP-2000.				
XX		XX				
PF		25-FEB-2000; 2000EP-0301439.				
XX		XX				
PR		25-FEB-1999; 99US-0121825.				
PR		05-MAR-1999; 99US-0123180.				
PR		09-MAR-1999; 99US-0123548.				
PR		23-MAR-1999; 99US-0125788.				
PR		23-MAR-1999; 99US-0126264.				
PR		29-MAR-1999; 99US-0126785.				
PR		01-APR-1999; 99US-0127462.				
PR		06-APR-1999; 99US-0128234.				
PR		08-APR-1999; 99US-0128714.				
PR		16-APR-1999; 99US-0129845.				
PR		19-APR-1999; 99US-0130077.				
PR		21-APR-1999; 99US-0130449.				
PR		23-APR-1999; 99US-0130510.				
PR		28-APR-1999; 99US-0130891.				
PR		30-APR-1999; 99US-0131449.				
PR		30-APR-1999; 99US-0132048.				
PR		04-MAY-1999; 99US-0132484.				
PR		05-MAY-1999; 99US-0132485.				
PR		06-MAY-1999; 99US-0132486.				

PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	28-JUL-1999;	99US-0145986.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	26-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	28-JUL-1999;	99US-0145986.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	26-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	28-JUL-1999;	99US-0145986.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	26-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159638.
PR	21-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.

Query Match 72.9%; Score 4209.5; DB 21; Length 959;

Best Local Similarity 80.4%; Pred. No. 0;

Matches 783; Conservative 80; Mismatches 92; Indels 19; Gaps 6;

Qy	111	MRSRMNTGGSGNVGHPKYDSGEIGLSKYDSGEIPRGVPSVTNSQ-MSGEIPGASPDHH	169
Db	1	MLGWHLTRGKGEEMGEPOYDK-----EVSHHHLPLTISRQDTSGEFSNASPERL	49
Qy	170	MMSPGTGNISRAPPY-VNHSPNPSREFSGSIGNVANKERVGKMKQDKGAIPMTNGTS	228
Db	50	SVSSTIAGGKRLPYSSDVNQSPNRRIVDPVGLGNVANKERVGKMKQDKNTGTV---ST	106
Qy	229	IAPSEGRAATDIDASTEYNMEDALLNDETROPLSRKYPFIASSKINPYRMVIVLRLVLSI	288
Db	107	QAASE-RGGVDIDASTDILADEALLNDEARQPLSRKYSIPSSRINPYRMVIMLRVLTLCL	165
Qy	289	FLHYRLTNPNVFNAYPLLLSVICEIFEALSILDPQFPKFPINRETYLDRLLALRYDREGE	348
Db	166	FLHYRITNPVFNALWLSVICEIFEALSILDPQFPKFPVNRETYLDRLLALRYDREGE	225
Qy	349	PSQLAAVDIFVSTVDPLKEPPITVANTVLSILAVDYPVKVSCYSDDGASMLTFDALAE	408
Db	226	PSQLAAVDIFVSTVDPLKEPLVTANTVLSILAVDYPVKVSCYSDDGAAMLSFSALAE	285
Qy	409	TSEFARKWVPVKKYDIEPRAPEYFCQKIDYLDKQVQPSFVKDRRAKREYEFKIRIN	468
Db	286	TSEFARKWVPFCKKYSTIEPRAPEYFAAKIDYLDKQVTSFVKDRRAKREYEFKIRIN	345
Qy	469	ALVSKALKVPEEGIMODGTPWPGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYSRE	528
Db	346	ALVSKALKCPEGWVMDGTPWPGNTRDHPGMIQVFLGQNGGLDAGNEPLRLVYSRE	405
Qy	529	KRPGFQHHKAGAMNALVRVSAYLTNGOYMLNLDGCDHYINNSKAVREAMCFMDPNLGPK	588
Db	406	KRPGFQHHKAGAMNALVRVSAYLTNGPFILNLDGCDHYINNSKALREAMCFMDPNLGKQ	465
Qy	589	VCYVQFPQRFQDIDRNDRYANRNTVFFDINLRGLDGTGGPVYVCTGCVFNRTALYGEPP	648
Db	466	VCYVQFPQRFQDIDRNDRYANRNTVFFDINLRGLDGTGGPVYVCTGCVFNRTALYGEPP	525
Qy	649	IKA--KKPGFLASICGGKKASKSKRSOKKSNKHWDSDVSPVFNLEDTEEGVEGAGFD	706
Db	526	IKVHKKPSLLSKLGGSRKKNKAKKESDKKSGRHTDSTVPVFNLDLDEEGVEGAGFD	585
Qy	707	DEKSVLMSQMSLEKRRFGQSAFVASTLMYEGVQPSSTPESLLKEAITHVSCGYEDKSEW	766
Db	586	DEKALLMSQMSLEKRRFGQSAFVASTLMENGVPPSPATPENLLKEAITHVSCGYEDKSDW	645
Qy	767	GTEIGWIYGSVTEIDILTGFKMHARGWSVVCMPKRPAPKGSAPINLSDRLNOVLRLWALGS	826
Db	646	GMEIGWIYGSVTEIDILTGFKMHARGWSIYCMPLKPAKGSAPINLSDRLNOVLRLWALGS	705
Qy	827	VEILFSRHCPWYGGGRKFLERFAYINTIYPLTSLPLIVYCIILPAICLLTGKFTMPE	886
Db	706	VEILFSRHCPWYGNGLKFLERFAYVNTIYPLTSLPLMYCTLPVCLFTNQFIIPQ	765
Qy	887	ISNLASITWFIATLFSIFATGILEMRWSGVGIDEWRNEQFWVIGGISAHLFAVFOGLLK	946





QY 1009 GPLFGKLFPAFWIVHLYPFLKGLMGQRNTPPTIVIVAVLLASIFSLWVRVDPF---T 1065  
 DB 1002 GPLFGKLFESVILHLYPFLKGLMGQRNTPPTIVIVWSILLASIFSLWVKIDPISPT 1061

QY 1066 TRLAGPNIQTGICNG 1080  
 DB 1062 QKAAA--LGQGVNC 1074

RESULT 12  
 RAY84112  
 ID AAY84112 standard; Protein; 1074 AA.  
 XX  
 AC AAY84112;  
 XX  
 DT 03-JUL-2000 (first entry)  
 XX  
 DE Amino acid sequence of a maize cellulose synthase.

KW Maize; cellulose synthase; stalk quality; stand; silage; cellulose;  
 KW transgenic plant; plant breeding marker.

XX Zea mays.

XX WO200009706-A2.

XX 24-FEB-2000.

XX 16-AUG-1999; 99WO-US18760.

XX 17-AUG-1998; 98US-0096822.

XX (PTON-) PIONEER HI-BRED INT INC.

PI Dhugga KS, Helentjaris TG, Bowen BA, Wang X;

XX WPI; 2000-224343/19.

XX N-PSDB; AAZ99506.

PT New genes which encode maize cellulose synthase polypeptides in plants  
 useful for modulating the expression of cellulose synthase in plants

XX and to produce transgenic plants expressing the novel protein -

PS Claim 15; Page 126-128; 119pp; English.

XX The present sequence represents a maize cellulose synthase polypeptide.  
 CC The cellulose synthase can be used for the improvement of stalk quality  
 CC for improved stand or silage. It also provides an increased concentration  
 CC of cellulose in the pericarp, hardening the kernel and improving its  
 CC handling ability. The sequences are used to produce transgenic plants  
 CC and seeds expressing the cellulose synthase. The polynucleotide is  
 CC used for modulating, preferably increasing, the level of the synthase  
 CC in a plant cell. The plants are preferably monocots. The polynucleotide  
 CC is also used as a probe or primer in the detection quantitation or  
 CC isolation of gene transcripts. The probes are useful in detecting  
 CC deficiencies in the level of mRNA in screenings for desired transgenic  
 CC plant, for detecting mutations in the gene, for monitoring upregulation  
 CC of expression or changes in enzyme activity in screening assays of  
 CC compounds, for detection of any number of allelic variants of the gene,  
 CC or for use as molecular markers in plant breeding programs. The  
 CC isolated nucleic acids of the present invention can also be used for  
 CC recombinant expression of their encoded polypeptides or for use as  
 CC immunogens in the preparation and/or screening of antibodies. The  
 CC proteins can be employed in assays for enzyme agonists or antagonists  
 CC of enzyme function or for use of immunogens or antigens to obtain  
 CC antibodies specifically immunoreactive with a protein.

XX Sequence 1074 AA;

Query Match 69.18; Score 3993.5; DB 21; Length 1074;  
 Best Local Similarity 68.4%; Pred. No. 0;  
 Matches 749; Conservative 134; Mismatches 153; Indels 59; Gaps 20;

QY 2 DGDADA-LKSGRHAGADVQCICADGLGTTLDGVDFTACDVCRFPCPCYEHKKECTQA 60  
 DB 23 DGDADVPKAPTKSANGQVCICGDTVGVSATGDFVACNECAFVPCPCYEHKKEGQC 82  
 QY 61 CLQCKTKYKRRHSGPAIRBEGDDTDADD-GSDFNYPASGTEDOKQKIADMRMSWRMNTG 119  
 DB 83 CPQCKTRYKROKGSPPRVHGGD-EEEDVDOLDNEFNY-----KQ----- 119  
 QY 120 GSGNVGHPKY----DSGEIGLS-KYDSGEIPRGVYVSVTN-SOMSGEIPGASPDH-MMS 172  
 DB 120 --GNGKGPFWOIQGDDADLSSARHD---PHHRIPRLTSGQSIGEIPDASPDHSIRS 173  
 QY 173 PTGN-ISRAPPYVYNHSPNPSREF-SGSIGNVAMKERVDMKMDKGAIPMTNGSIA 230  
 DB 174 PTSSYVDPSPVPV--RIVDPKDLNSYGLNSVDKERVESWRVKQDKNMQVTN----- 226  
 QY 231 PSEGRAATDIDASTEYNMEDALLNDETQPLSRKVPITASSKINPYRMVIVLVLVLSIFL 290  
 DB 227 -KYPEARGDME-GTGSNGEDQMVDARLPLSRIVPISSNOLNDYRIVILRLILICFFF 284  
 QY 291 HYLTPNVRNAYPLWLLSVICEIWFALSWILDQFPKFPINRETYLDRALRYDREGEPS 350  
 DB 285 QYRISHPVRNAYGLWLVSVICEVWFALSWILDQFPKFPINRETYLDRALRYDREGEPS 344  
 QY 351 QLAADVDFVSTVDPLKEPPIVTANTVLSILAVDYPVKVSYSDGASMLTDALAEYS 410  
 DB 345 QLAIDVFVSTVDPLKEPPLITANTVLSILAVDYPVKVSYSDGASMLTFESLSETA 404  
 QY 411 EFARKWVPFKYKDYIEPRAPEFYFCQKIDYLDKQVQSFVKDRRAMKREVEEKKIRINAL 470  
 DB 405 EFARKWVPFKYKDYIEPRAPEFYFCQKIDYLDKQVQSFVKDRRAMKREVEEKKIRINAL 464  
 QY 471 VSKALKVPEEGKIMODGTPWPGNNTDRHPGMIQVFLGHSGGLDTEGNEHPLRYVVSREKR 530  
 DB 465 VAKAQKVPPEGTWTADGTAWPGNPNRHPGMIQVFLGHSGGLDTEGNEHPLRYVVSREKR 524  
 QY 531 PGFQHHKAGAMNALVRYSAVLTNGOYMLNDCDHYLNNSKAVREACMFMDNPLGPOVC 590  
 DB 525 PGFQHHKAGAMNALVRYSAVLTNGAYLLNVDCDHYFNSSKALREACMFMDPALGRKTC 584  
 QY 591 VYVQFPQFEDGIDRNDRYANRNTVFFDINLRGLDGIQGVYVGTGCVFNRTAIYGEYPIK 650  
 DB 585 VYVQFPQFEDGIDLHDYRANRNVFFDINMKGLDGIQGVYVGTGCVFNRTAIYGEYPIK 644  
 QY 651 AK--KPGFLASLCGSKKASKSKRSDDKKSNKHVDSVPVFNLEIDIEGVEGAGDDE 708  
 DB 645 EADLEPNIVKSCGRRK-RANKSYMDSQSRIMARTESSAPFNWEDIEGIE--GYEDE 701  
 QY 709 KSVLMSQSLKRRFGQSAFVASTIMEYGVQSPSTPESLLKEAHIVISCGYEDKSEWGT 768  
 DB 702 RSVLMSQSLKRRFGQSPFIFASTFTQGGIPPTNPNASLLKEAHIVISCGYEDKSEWGT 761  
 QY 769 EIGWYIGSVTEDLTGFKMHARGWRSVYCMKRPAPKAPINLSDRNLNQLRWALGSVE 828  
 DB 762 EIGWYIGSVTEDLTGFKMHARGWRSVYCMKRPAPKAPINLSDRNLNQLRWALGSVE 821  
 QY 829 ILFSRHCPWYGVYGRGLKFLERFAYINTIYPLSLPLVYVYILPAICLTGKFMPEIS 888  
 DB 822 ILLSRHCPTWYGVYGRGLKFLERFAYINTIYPLSLPLVYVYILPAICLTGKFMPEIS 881  
 QY 889 NIASWTFALTFLSFATGILEMRWSGVGIDEMWRNEQFWIIGGISAHLEAVFGGLKLVLA 948  
 DB 882 NYAGFFILLFASIFATGILEMRWSGVGIDEMWRNEQFWIIGGISAHLEAVFGGLKLVLA 941  
 QY 949 GIDTNTVTSKANDSEGDFAEIYMFKWTLLIPPTTILINMVGVAAGTSYAINSGYSQSW 1008  
 DB 942 GIDTNTVTSKANDSEGDFAEIYMFKWTLLIPPTTILINMVGVAAGTSYAINSGYSQSW 1001  
 QY 1009 GPLFGKLFPAFWIVHLYPFLKGLMGQRNTPPTIVIVAVLLASIFSLWVRVDPF---T 1065  
 DB 1002 GPLFGKLFESVILHLYPFLKGLMGQRNTPPTIVIVWSILLASIFSLWVKIDPISPT 1061







---







```

QY 436 QKIDYLDKQVPSFVKDRAMKREYEEFKIRINALYSKALKYPERGWINQDTPPGNNT 495
DB 419 AKIDYLDKQVTSFVKDRAMKREYEEFKIRINALYSKALKYPERGWINQDTPPGNNT 478
QY 496 RDHPGMIQVFLHSGGLDTGEGNELPRLVVSREKRGPGFQHHKKAGAMNALVRVSAVLTNG 555
DB 479 GDHPGMIQVFLGQGLDAEGNELPRLVVSREKRGPGFQHHKKAGAMNALVRVSAVLTNG 538
QY 556 OYMLNDCDHYINNSKAVREACMLDPNLPQOVYVQPPORFDGIDRNDRYANRNTVFF 615
DB 539 PFLNDCDHYINNSKALREACMLDPNLPQOVYVQPPORFDGIDRNDRYANRNTVFF 598
QY 616 DINLRGLDGIQGVVYVGTGCVFNRTAIYGEYPIKA--KKPGFLASLCGGKKKSKSKR 673
DB 599 DINLRGLDGIQGVVYVGTGCVFNRTAIYGEYPIKVKHKKPSLLSKLCSGSRKKSKAKK 658
QY 674 SSDKKSNKHVDSSVPVFNLEDIEEGVEGAGDEDEKSVLSMSLEKRFQSGAAVASTL 733
DB 659 ESDKKSGRHTDSTVPVFNLEDIEEGVEGAGDEDEKALMSLEKRFQSGAVASTL 718
QY 734 MEYGGVQPSQSTPSLLEKAIHVISCYEDKSEWGTGEGWYGSVTEIITGFKMHARGWR 793
DB 719 MENGVPSPSATPENLLEKAIHVISCYEDKSDWMEIGWYGSVTEIITGFKMHARGWR 778
QY 794 SVTCMKRPAFKSAPINLSRDLNQLRWALGSVELFSRHCPWYGYGRKFLERFAY 853
DB 779 SIYCMPLKPAFKSAPINLSRDLNQLRWALGSVELFSRHCPWYGYGRKFLERFAY 838
QY 854 INTIYPLTSLPVLVYCIPLAICLTGKFIPEISNLASIFATLFLSIFATGILEMRWS 913
DB 839 VNTIYPTISPLMYCTLLAVCLFTNQFIIPQISNIASIFLFLSIFATGILEMRWS 898
QY 914 GVGIDEMWNEQFWVIGGISAHLFAVFOGLLKVLAGIDTNFTVTSKANDEEDGFAELMYF 973
DB 899 GVGIDEMWNEQFWVIGGISAHLFAVFOGLLKVLAGIDTNFTVTSKANDEEDGFAELMYF 958
QY 974 KWTLLIPTTILINMVGWAGTSYAINSGYQSMGPLEFGKFFFAFWVIVHLYPFLKGLM 1033
DB 959 KWTLLIPTTILINMVGWAGTSYAINSGYQSMGPLEFGKFFFAFWVIVHLYPFLKGLM 1018
QY 1034 GRONRPTIYVWVLLASIFLSLLWVRDPFTTLAGRNIOFCGNC 1080
DB 1019 GRONRPTIYVWVLLASIFLSLLWVRDPFTTLAGRNIOFCGNC 1065

RESULT 2
T05351
cellulose synthase (EC 2.4.1.-) catalytic chain RSW1 - Arabidopsis thaliana
N:Alternate names: protein F8B4.110
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T05351
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De
  ewes, H.W.; Mayer, K.F.X.; Schueller, C.
  submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15409
A:Accession: T05351
A:Molecule type: DNA
A:Residues: 1-1081 <BEV>
A:Cross-references: EMBL:AL034567
A:Experimental source: cultivar Columbia; BAC clone F8B4
C:Genetics:
A:Gene: RSW1
A:Map position: 4
A:Intons: 27/3; 93/1; 150/3; 189/1; 252/2; 341/2; 456/3; 502/3; 544/3; 615/3; 704/1; 76
A:Note: F8B4.110
C:Keywords: glycosyltransferase; hexosyltransferase; P-loop

Query Match 68.6%; Score 3963.5; DB 2; Length 1081;
Best Local Similarity 68.0%; Pred. No. 3.1e-282;
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;

```

```

QY 2 DGDADALSGRHGAGDYCOICA 1
DB 25 DGGTKPLKMN--GQICQICG 81
QY 62 LQCKTKYKRRHGRGSPAIRGEGD 121
DB 82 POCKTRFRHRHGRSPRVEGDEDE 133
QY 122 GNVGHPKYDGEIGLSKYD-- 178
DB 134 ----SSRHEOSIPLLTHGHTW 184
QY 179 RRAPFPVYVNHSPNREF--SGS 237
DB 185 PRQPVV--RIVDPKDLNSYG 238
QY 238 TDIDASTEYNMEDALLNDETRO 297
DB 239 -EIE--GTGSGNEELQMADDTRI 296
QY 298 VRNAPYMLLSVICIWFALS 357
DB 297 VKNAPYMLLSVICIWFALS 356
QY 358 FVSTVDPKPEPIVYANTVLS 417
DB 357 FVSTVDPKPEPIVYANTVLS 416
QY 418 PFVKYDIETPRAPFIFYCOKIE 477
DB 417 PFCKKFNIEPRAPFIFYAOKI 476
QY 478 PEEGWIMQDGPWPGNNTROH 537
DB 477 PEEGWIMQDGPWPGNNTROH 536
QY 538 KAGAMNALVRYSVAVLTNGQYM 597
DB 537 KAGAMNALVRYSVAVLTNGAYL 596
QY 598 FDGIDRNDRYANRNVFEDIN 655
DB 597 FDGIDRNDRYANRNVFEDIN 656
QY 656 FLA--SLCGKKKKSKSKRKS 714
DB 657 IIVKSCCGSRKKKSKSKYNY 714
QY 715 QMSLEKRFQSGAAFAVASTLME 774
DB 715 QRSVEKRFQSGSPVFIATFME 774
QY 775 GSVTEIITGFKMHARGWR 834
DB 775 GSVTEIITGFKMHARGWR 834
QY 835 CPLWYGYGRKFLERFAYIN 894
DB 835 CPIWYGYGRKFLERFAYIN 894
QY 895 FIALFLSIFATGILEMRWS 954
DB 895 FILLFISIAVTGILELRWS 954
QY 955 TVTSKANDEEDGFAELMYF 1014
DB 955 TVTSKANDEEDGFAELMYF 1014
QY 1015 LFFAFWVIVHLYPFLKGLMS 1072
DB 1015 LFFALWVIAHLYPFLKGLLS 1070

```



Qy 348 EPSQLAAVDIEFVSTVDPLKEPIPIVANTVLSILAVDYPVDKVCYVSDGASMLTFDALA 407  
Db 334 EPSQLAPVDVFSVDPKMEPLPVTANTVLSILAVDYPVDKVCYVSDGASMLTFEALS 393  
Qy 408 ETSEFARKWPFVKKYDIEPRAPPEYFCQKIDYLDKVKQPSFVKDRAMKREYEEFKIRI 467  
Db 394 ETAFESKWWPFCKFNIEPRAPPEYFSQKIDYLDKIKIOPSFKERRAMKREYEEFKVRI 453  
Qy 468 NALYSKALKVPEEGWIMQDGTWPFGNTRDPHGMIOVFLGSHSGGLDTGEGNELPLVYVSR 527  
Db 454 NILVAKAQKIDEDGWTMEDGTSWFGNPRDPHGMIOVFLGSHSGGLDTGEGNELPLVYVSR 513  
Qy 528 EKRPFGFHHKAGAMNALVRVSAVLNTQYMLNDCDHYINNKAIVREAMCFMDPNLGP 587  
Db 514 EKRPFGFHHKAGAMNALIRVSAVLNTGAYLLNVDCHYFNNSKAIRKAMCFMDPDAIGK 573  
Qy 588 QVCYVQRFQDFGIDRNDRYANRNTVFDINLRGLDGGIQGVYVGTGCVFNRTALYGEPE 647  
Db 574 KCCYVQRFQDFGIDLHRYANRNTVFDINLRGLDGGIQGVYVGTGCVFNRTALYGEPE 633  
Qy 648 PIKAK--KPGFLASLCGGKKKASKSK--RSSDKKKSNKHVDSSVPVFNLEDIEBEGVAG 704  
Db 634 VLTEEDLEPNIIVKSCFGSRKKGSKKIPNYEDNRSKRSDSNVPLFNMEDIDEDVE--G 691  
Qy 705 FDDEKSVLMSOMLEKRFQGSAAVASTLMEYGGVPOSSTPESLLKKAHIVISGIEDKS 764  
Db 692 YEDEMSLVSKRLEKRFQGSVFIATFMEOGGLPSTNTPLTLKKAHIVISGIEYAKT 751  
Qy 765 EMGTEIGWYGSVTEIDLTGFKMHARGWRSVYCMKPAFKGAPINLSDRNLQVLRWAL 824  
Db 752 DWGKEIGWYGSVTEIDLTGFKMHARGWISVYCPSPAFKGSAPINLSDRNLQVLRWAL 811  
Qy 825 GSVTEILSRHCPWYGVGRUKLEFPAYINTTIVPLTSLVLYCILPATCLTGTGKFM 884  
Db 812 GSIEILLSRHCPWYGVGRUKLERIAYINTTIVPITSIPILAYCMPLPACLTNTFII 871  
Qy 885 PEISNLASIWTFILFSLFATGILEMRWSGVGIDEMWNEQFVWIGGISAHLFVFOGLL 944  
Db 872 PEISNLASLCFMLFASIAAILELWSDVALEDWNEQFVWIGGISAHLFVFOGLL 931  
Qy 945 KVLADIDNFTVTSKANDEEDGFAELVFKWTTLLIPPTTILINMVGVAAGTSVAINSG 1004  
Db 932 KVFAGIDNFTVTSKANDEEDGFAELVFKWTTLLIPPTTILINMVGVAAGTSVAINSG 991  
Qy 1005 YQSWGPLEGKLFPAFWIVHLVPLFLKGLMQRNRTPTTIVVAVLLASIFSLMWVRDPPF 1064  
Db 992 YQSWGPLMGKLLFAFWVVAHLYPFLKGLLQRNRTPTTIVVAVLLASIFSLMWVRINPF 1051  
Qy 1065 TTRLAGPNIQTCGI 1078  
Db 1052 -----VSTGV 1057

RESULT 5  
T08583  
cellulose synthase (EC 2.4.1.-) catalytic chain - Arabidopsis thaliana  
N:Alternate names: protein f22f8.250  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 22-Oct-1999  
C:Accession: T08583; T09014  
R:Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; May  
submitted to the Protein Sequence Database, May 1999  
A:Reference number: 216442  
A:Accession: T08583  
A:Molecule type: DNA  
A:Residues: 1-1084 <BEV>  
A:Cross-references: EMBL:AL050351; GSPDB:CN00062; ATSP:T22F8.250  
A:Experimental source: cultivar Columbia; BAC clone T22f8  
R:Arioli, T.; Peng, L.; Betzner, A.S.; Burn, J.; Wittke, W.; Herth, W.; Camilleri, C.; H  
Science 279, 717-720, 1998  
A:title: Molecular analysis of cellulose biosynthesis in Arabidopsis.  
A:Reference number: Z13745; M0UD:98111412; PMID:9445479

A:Accession: T09014  
A>Status: translated from GB/EMBL  
A:Molecule type: mRNA  
A:Residues: 1-1084 <CAR>  
A:Cross-references: EMBL:AF02717  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Gene: ATSP:T22F8.250; Ath-A  
A:Map position: 4  
A:Introns: 27/3; 156/3; 191/1; 2  
C:Keywords: cell wall synthesis;  
Query Match 63.5%;  
Best Local Similarity 62.4%;  
Matches 687; Conservative 1  
Qy 1 MDGADADA-LKSGRHGAGDVCOI 1  
Db 20 INADESARIRSVQELSQTCOI 1  
Qy 60 ACLOCKTKYKRRGSPAIRGEE 1  
Db 80 ACPOCKTRYKRIKGSPRV--- 1  
Qy 120 GSGNVGHPKYDSGIGLSKYDS 1  
Db 136 RGLDLSAP----- 1  
Qy 178 SRRAPFYVNHSPNPS----- 1  
Db 177 NRVYPAPFTDSSAPPQARSVY 1  
Qy 227 TSIAPSGGRATDIDASTENY 1  
Db 235 ---GNNGRGSNDDD---ELDE 1  
Qy 287 SIFHYRLNTPVRNAYPIWLLS 1  
Db 288 GLFPHYRLHPVNDAYGLWLT 1  
Qy 347 GEPQLAAVDIFVSTVDPLKE 1  
Db 348 GKPSGLAPDVFEVSTVDPLKE 1  
Qy 407 AETSEFARKWPFVKKYDIEPR 1  
Db 408 SDTAEFARKWPFCKFNIEP 1  
Qy 467 INALVSKALKVPEEGWIMQDG 1  
Db 468 INALVATAQKVPPEGWIMQDG 1  
Qy 527 REKRPFGFHHKAGAMNALVR 1  
Db 528 REKRPFGFHHKAGAMNSLIR 1  
Qy 587 PQVCYVQRFQDFGIDRNDRY 1  
Db 588 KVCYVQRFQDFGIDRNDRY 1  
Qy 647 PPIKAKPGFLAS-----L 1  
Db 648 APKKKKPGKTCNCWPKWCC 1  
Qy 700 VEGAGFDDERSVLMQSOMLEK 1  
Db 704 VIVPVSNNVEKRESEATQLEK 1  
Qy 760 YEDKSEWTEIGWYGSVTEID 1  
Db 764 YEDKTEWKEIGWYGSVTEID 1  
Qy 820 LRWALGSVEILFNRHCPWYGV 1

J  
ID:g2827140; PIDN:AAC39335.1; PID:g2827141  
bia  
343/2; 458/3; 504/3; 546/3; 617/3; 708/1; 773/3  
osyltransferase; hexosyltransferase  
are 3666.5; DB 2; Length 1084;  
ad. No. 1.9e-260; Indels 59; Gaps 16;  
Mismatches 191;  
GTTLDGVDFTACDVCRFPVPCRCYEHKREGTQ 59  
ELTVSSSELFVACNECAPVCRPCYERREGNQ 79  
SADGSDFNYPASGTEDOKQKIADMRMRNTG 119  
SEEDIDDDLEYEFDHGM--PEHAACAALSSRLNTG 135  
RGYPVSVTNSQMSGEIPGASPDHMMSP--TGN 177  
SSQIPLLTYCEDADAM--YSRHALIVPSTGYG 176  
REFSGSIGNVAKKERVDCGKWKQ-DKGAIPWTNG 226  
KEY--GYGSAVAKORMEVKKRQSEKLOVHKEG 234  
LNDETROPLSRKVPIASSKINPYRMVIVLRVLV 286  
WMDEGROPLSRKLPTRSSRINPYRMILCLRIL 287  
LWFALSWILDQFPKWPFPINRTYLDRLALRYDRE 346  
LWFVAVSWILDQFPKWPYERETYLDRLSRLEKE 347  
ANTVLSILAVDYPVDKVCYVSDGASMLTFDAL 406  
ANTVLSILAVDYPVDKVCYVSDGASMLTFEAL 407  
YFCQKIDYLDKVKQPSFVKDRAMKREYEEFKIR 466  
YFSQKMDYLNKRVHPAFVRRRRAMKRDYEEFKV 467  
ANTRDHPGMIOVFLGSHSGGLDTGEGNELPRLVYVS 526  
ANNVRDHPGMIOVFLGSHSGVTRDGTGNEPRLVYVS 527  
INGQYMLNDCDHYINNKAIVREAMCFMDPNL 586  
SNAPYLLNVDCDHYINNKAIVRESCFMDPQSG 587  
VFFDINLRGLDGIQGVYVGTGCVFNRTALYGYE 646  
VFFDINLRGLDGIQGVYVGTGCVFNRTALYGYE 647  
VFFDINLRGLDGIQGVYVGTGCVFNRTALYGYE 647  
KASKSKRSDSKKSNKHVDSSVPVFNLEDIEEG 699  
K--KSKTKAKDKKTK--ETSKQIHALENVDEG 703  
AAFAVASTLMEYGGVPOSSTPESLLKKAHIVIS 759  
PVFVASAVLQNGVPRNAPSACLLREALQVLSG 763  
KMHARGWRSVYCMKPAFKGAPINLSDRNLQV 819  
KMHCHGWRVYCMKPAFKGAPINLSDRNLQV 823  
KFLERFAINTTIVPLTSLVLYCILPATCLLT 879

```
||||| 824 LRWALGSVEIFLSRHCPITWYGGGLKWLRFYSYINSVVPWTSPLIVVCSLPVCLLT 883
QY 880 GKETMPEISNLASITWFLALFSLIPATGILEMRSGVGDWNRNEOFWVIGGISAHLFAV 939
Db 884 GRFIVPEISNAGLFLDMFLISTIAVTGILEMQGGVGDWNRNEOFWVIGGASSHLFAL 943
QY 940 FOGLKVLGADTNTFTVTSKANDREGDAELMYFKWTTLLIPPTTLIIINMVGVVAGTSY 999
Db 944 FOGLKVLGAVNTFTVTSKAA--DGAFFSELYIFKWTLLIPPTTLIIINIGVGVSD 1002
QY 1000 AINSYOSMGPLFGKLEFAFVIVHLYPFLKGLMGRONRPTTIVVAVLLASIFSLWV 1059
Db 1003 AINSYDWMGPLFGRLFFALVIVHLYPFLKGLMGKDKMPTTIVVMSILLASILLWV 1062
QY 1060 RYDPTTTLRAGPNITQTCGNC 1080
Db 1063 RVNPFVAK--GGPVLEICGLNC 1082

RESULT 6
T52028
cellulose synthase [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52028
R:Joshi, C.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z25890
A:Accession: T52028
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1081 <TOS>
A:Cross-references: EMBL:AF062485; PIDN:AAC29067.1

Query Match 62.0%; Score 3583; DB 2; Length 1081;
Best Local Similarity 61.1%; Pred. No. 2.6e-254;
Matches 676; Conservative 162; Mismatches 198; Indels 72; Gaps 15;

QY 1 MDGADADA-LKSGRHGADVCOICADGLGTLTLDGDTFTACDVCRCPCYEHREGTQ 59
Db 17 INADENARISVQELSGTQCIDEIETLVDGEPFVACNECAPPCVRCPCYERREGNQ 76

QY 60 ACLOCKTKYKRRHSGPAIRREGDDTDADGSDFNYPASGTE-DQOKIADRMRSWRMNT 118
Db 77 ACPOCKTRFKLKGSPRVEGDEEDDIDLDNEPEYGNNGIGFDQ--VSEGMSISRRNS 133

QY 119 GSGNVGHPKYDGSGEIGLSKYDSCGEIPRGVYPSVTNSQMSGEIPGASPDHH----- 169
Db 134 -----GFPQSD-----LDSAP-PGSQLPLTYGDEVEI---SSDRHALIVPPSLG 175

QY 170 -----MMSPTGNISRRAPFPYYVNHSPNPSREFSGSIGNVANKERVVDGKMKQDKG 219
Db 176 GHGNRVHPVLSLDPFVAHRLMLVPQDKLAVY-----GYGVSANKRMEWKRQNE- 227

QY 220 AIPWTNGTSIAPSGRAATDIDASTEYNNMEDALLNDETROPLSKRPVIASSKNPVRMVI 279
Db 228 -----KIQVVRHGGDPDFDGDADPFM-----MDGROPGLSMKIPKISKNIPYRMLI 276

QY 280 VLRLVLSITFLYRLTNVRNAYPLWLLSVICEIWFALSVICELTDOFPKWFFPINRETYDLRL 339
Db 277 VLRLVILGLFFHYRLHPVKDAYALWLLSVICEIWFALSVVLDQFPKWYPIERETYLRL 336

QY 340 ALRYDREGEPQLAANDIFVSTVDPLKEPPIVTANTVLSILAVDYPVDKVCSTVSDGAS 399
Db 337 SLRYEKEGKPSGLSPVDVFTVDPLKEPPLITANTVLSILAVDYPVDKVCSTVSDGAA 396

QY 400 MLTFDALAETSEFAKWPVKYKIDYKDKVQPSVVKDORRAMKRE 459
Db 397 MLTFDALAETAEFAKWPVKCKYIEPAPWYFCHKMDYLNKNVHPAVERRAMKRD 456

QY 460 YEEFKIRINALVSKALVPREGTMDQGTWPNGNTRDHPGMIOVFLGSGGDLDTGREL 519
||||| 824 LRWALGSVEIFLSRHCPITWYGGGLKWLRFYSYINSVVPWTSPLIVVCSLPVCLLT 883
QY 880 GKETMPEISNLASITWFLALFSLIPATGILEMRSGVGDWNRNEOFWVIGGISAHLFAV 939
Db 884 GRFIVPEISNAGLFLDMFLISTIAVTGILEMQGGVGDWNRNEOFWVIGGASSHLFAL 943
QY 940 FOGLKVLGADTNTFTVTSKANDREGDAELMYFKWTTLLIPPTTLIIINMVGVVAGTSY 999
Db 944 FOGLKVLGAVNTFTVTSKAA--DGAFFSELYIFKWTLLIPPTTLIIINIGVGVSD 1002
QY 1000 AINSYOSMGPLFGKLEFAFVIVHLYPFLKGLMGRONRPTTIVVAVLLASIFSLWV 1059
Db 1003 AINSYDWMGPLFGRLFFALVIVHLYPFLKGLMGKDKMPTTIVVMSILLASILLWV 1062
QY 1060 RYDPTTTLRAGPNITQTCGNC 1080
Db 1063 RVNPFVAK--GGPVLEICGLNC 1082

RESULT 6
T52028
cellulose synthase [imported] - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
C:Accession: T52028
R:Joshi, C.
submitted to the EMBL Data Library, May 1998
A:Reference number: Z25890
A:Accession: T52028
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1081 <TOS>
A:Cross-references: EMBL:AF062485; PIDN:AAC29067.1
```

```
Db 457 YEEFKIRINALVATAQVPEDGWTMDGTWPNGNSVRDHPGMIQVFLGSGVDRVENNEL 516
QY 520 PRLVYVREKRPFGQHHKAGAMNALVRVSAVLNTNGOYMLNDCDHYINNSKAVRAMCF 579
Db 517 PRLVYVREKRPFGDHHKAGAMNSLIRVSGVLSNAPYLLNVDCHYINNSKALRAMCF 576
QY 580 LMDPNLGPQVCYVOPFORFDGIDNRNRYANRTVFFDINLRGDLGDIQGPVYVGTGCVFNR 639
Db 577 NMDPQSGKKICYVOPFORFDGIDNRNRYANRTVFFDINLRGDLGDIQGPVYVGTGCVFNR 636
QY 640 TAIYGYEPPKAKKPGFLAS-----LCGKKKASKSKRSDKKSKNKHVDSSVPVEN 692
Db 637 QALYGFDAKPKKKGPRKTCNCWKWLLCFGSRKNRKAKTVAADKKKNR--PASKQIHA 694
QY 693 LEDIEEGVEGAGFDDEKSVLMSQMSLEKRFQSQAAFAVSTLMEYGGVPOSSTPESLLKEA 752
Db 695 LENIEEGRHGKVLNVEQSTEAMQMLQKYQSPVFVASARLENGMGGMARNASAPACLLKEA 754
QY 753 IHVISCYVEDKSEWGTETGWTYGSVTEDILTGFKMHARGWRSVYVCMKPAFKGSAFINL 812
Db 755 IOVISRGYEDKTEWKGKEIGWYGSVTEDILTGSKMHSHGWRHVCTPKLAAFKGSAPINL 814
QY 813 SDRNLQNLRWALGSVEILFSRHCPITWYGGGLKWLRFYSYINSVVPWTSPLIVVCSYL 872
Db 815 SDRHLQVLRWALGSVEILFSRHCPITWYGGGLKWLRFYSYINSVVPWTSPLIVVCSL 874
QY 873 PAICLLTGKFTIMPEISNLASITWFLALFSLIPATGILEMRSGVGDWNRNEOFWVIGGI 932
Db 875 PAICLLTGKFTIVPEISNAYSAILEMAFSSAITGILEMQGKVGIDWNRNEOFWVIGV 934
QY 933 SAHLFAVFOGLKVLGADTNTFTVTSKANDREGDAELMYFKWTTLLIPPTTLIIINMVG 992
Db 935 SAHLFALFOGLKVLGADTNTFTVTSKAA--DGEFSDLYLFKWTSLIPPTTLIIINVG 993
QY 993 VVAGTSYAINSGYOSMGPLFGKLEFAFVIVHLYPFLKGLMGRONRPTTIVVAVLLAS 1052
Db 994 VIVGVSDAISNGYDSMGPLFGRLFFALVIVHLYPFLKGLLGLGKODRPTTIVVMSILLAS 1053
QY 1053 IFSLLWRVDPFTTLRAGPNITQTCGNC 1080
Db 1054 ILTLWVRVNFVAK--GGPVLEICGLNC 1080

RESULT 7
H84604
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: H84604
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: H84604
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1088 <STO>
A:Cross-references: GB:AE002093; NID:g4417271; PIDN:AAD20396.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g21770
A:Map position: 2

Query Match 61.9%; Score 3579; DB 2; Length 1088;
Best Local Similarity 61.1%; Pred. No. 5.1e-254;
Matches 677; Conservative 156; Mismatches 199; Indels 76; Gaps 20;

QY 4 DADALKSGRHGADVCOICADGLGTLTLDGDTFTACDVCRCPCYEHREGTQACLO 63
Db 24 DTARIRSAEELSGQTKCRDEIETDNGEPTIACNECAPPCVRCPCYERREGNQACPO 83
QY 64 CKTKYKRRHSGPAIRREGDDTDADGSDFNYPASGTE-DQOKIADRMRSWRMNT 117
```

Db 84 CGTRYKRIKSPRVGEDEEDD-DIDDLEHEFYGMD---PEHVTE-----AALYYMRLN 132  
Qy 118 TG-GSGNVGHPKYDSGEIGLSKYDSGEIPRGYVPSVNTSOMSGEIPGASPDHNMSP--- 173  
Db 133 TGRGTDEYSH-----LYSAS--PGSEVPLTYCDESDM--YSDRHAIIVPST 177  
Qy 174 -TGNISRRAPF-----PYNVHSPNPSREFS-GSIGNVANKERVGVGKMKO-DKGA 220  
Db 178 GLGNRVHVHPETDPSFASHTPRMV-----PQKDLTVYGYGVAVKDRMEVWKQIQIEKLQ 232  
Qy 221 IPTMNGTSIAPSEGRAADTDASTEYNMEDAL-LNDETROPILSRKVPITASSKINPYRMI 279  
Db 233 VVKNERVNDGGDGFIVDELQ-----DPGLPMWDEGRQPLSRKLPRSRIRNIPYRMLI 285  
Qy 280 VRLVVLVLIHYLHRYLNTVNRNAYPLWLVSICEIWFALSWILDOPFKWFPINRETYLDRL 339  
Db 286 FCRLAILGLFFHYRLHPVNDFAFGLWLTSTVCEIWFVAVSWILDQFPKWPYTERETYDLRL 345  
Qy 340 ALRYDREGEPSOLAAVDIFVSTVDPLEKPPITVTANTVLSILAVDPVDKVSVCYVSDGAS 399  
Db 346 SLRYEKSKPSELAPVDVFTVDPLEKPPITVTANTVLSILAVDPYVEKVACVYSDDGAA 405  
Qy 400 MLTFDALAETSEFARKWVPFKYKDYIEPRAPFYPFCQKIDYLDKQVQPSFYKDRAMKRE 459  
Db 406 MLTFEALSYTAEFARKWVPFKYKDYIEPRAPFYPFCQKIDYLDKQVQPSFYKDRAMKRD 465  
Qy 460 YEEFKIRINALVSKALKVPEBGWIMODGTPWPGNTRDPHGMIOVFLGHSGGLDTEGEL 519  
Db 466 YEEFKVKNALVSKVQKPEDGWTMQDGTWPGNVRDPHGMIOVFLGHSGVCDMDGNEI 525  
Qy 520 PLRYVVSREKRPFGQHHKAGAMNALRVSAVLNTGOYMLNDCDHYINNASKAVREAMCF 579  
Db 526 PLRYVVSREKRPFGQHHKAGAMNLSIRSAVLSNAPYLLNVDCDHYINNASKAVREAMCF 585  
Qy 580 LMDPNLGPQVCYVOPQRFEDGIDRNDRYANRNTVFEDINLGLDGIQGPVYVGTGCVFNR 639  
Db 586 MMDPOSGRKICVYVOPQRFEDGIDRNDRYANRNTVFEDINLGLDGIQGPVYVGTGCVFNR 645  
Qy 640 TAIYGEYEPPIKAKPGFLAS-----LCGGKKKASKSKRSDDKKSNKKNVDSSVFPN 692  
Db 646 QALYGFADAPKPKPPGRTCNCPKWCCLCCGMKKKTKGKVDNQRKPK---ETSKQIHA 702  
Qy 693 LEDIEEGVAGFDDKESVLMSONSLEKRFQGSAAVASTLMYEGGVQPSSTPESLLKEA 752  
Db 703 LEHIEEGQVYNAENNSSET--AQLKLEKKFGQSPVLVASTLLLLNGGVPSNVNPAASLLRES 760  
Qy 753 IHVISCYGEDKSEWTEIGWYGVSTEDILTFGKMHARGWRSVYCMKRPAPKAGSAPINL 812  
Db 761 IQVISCYGEERTWEGKEIGWYGVSTEDILTFGKMHARGWRSVYCMKRPAPKAGSAPINL 820  
Qy 813 SDRNLQVLRWALGSVEILFHSRCHPLWYGYGGRKFLERFAINTIYPLTSLPLLVYCIL 872  
Db 821 SDRLQVLRWALGSVEILFHSRCHPLWYGYGGRKFLERFAINTIYPLTSLPLLVYCIL 880  
Qy 873 PAICLLTCKFTMPETSINLASWFLATFLSIFATGILEMRWSVGIDEWNRNEQFWLTVGGI 932  
Db 881 PAICLLTCKFTMPETSINLASWFLATFLSIFATGILEMRWSVGIDEWNRNEQFWLTVGGI 940  
Qy 933 SAHLFAVFGQLLKVLAGIDTFTVTSKANDEEGDFAELMYFKWTLLIPPTTILINMVG 992  
Db 941 SSSLFALPQGLLKVLAGVSTNFTVTSKAAD--DGEFSELYIFKWTSLIPPTTILINIVG 999  
Qy 993 VVAGTSYAINSGYOSWGLPFLCKLFAFWIVHLYPFLKGLMGGRNQRPPTIIVWAVLLAS 1052  
Db 1000 VIVGVSDAINNGYOSWGLPFLCKLFAFWIVHLYPFLKGLMGGRNQRPPTIIVWAVLLAS 1059  
Qy 1053 IFSLWVRVDPFTTRLAGPNITQTCGNC 1080  
Db 1060 ILLTLWVRVNFVSK-DGPVLEICGLDC 1086

RESULT 8

Tl10797  
cellulose synthase (EC 2.4.1.-)  
C:Species: Gossypium hirsutum (upland cotton)  
C:Date: 16-Jul-1999 #sequence\_re  
C:Accession: Tl10797  
R:Pear, J.R.; Kawagoe, Y.; Schre  
Proc. Natl. Acad. Sci. U.S.A. 93  
A:Title: Higher plants contain h  
A:Reference number: Z1152; MUI  
A:Accession: Tl10797  
A:Status: preliminary; translate  
A:Molecule type: mRNA  
A:Residues: 1-974 <PEA>  
A:Cross-references: EMBL:U58283;  
A:Experimental source: strain A  
C:Genetics:  
C:Function:  
A:Gene: celA1  
A:Description: involved in the s  
C:Keywords: glycosyltransferase;  
Query Match 60.2%;  
Best Local Similarity 61.5%;  
Matches 657; Conservative  
Qy 18 VCQICADGLGTTLDGDVFTACD  
Db 8 VCHTCGEHVGNGVEPFVACH  
Qy 78 RGEEDGTDADDGSDFNYPASG  
Db 59 -----DENLLDD-----VEK  
Qy 137 SKYDSGEPGRGYVPSVTNSQMS  
Db 97 STLDS-----EM 103  
Qy 197 SGSTGNVAMKERVGDWKMQRD  
Db 104 AEDNGSIWNRVSWKREKNR  
Qy 257 TROPILSRKVPITASSKINPYRM  
Db 155 S-QPLSLIIPKSRAPYRTV  
Qy 317 LSWILDQPKWFPINRETYLDE  
Db 214 FSWILDQPKWYVNRRETYID  
Qy 377 LSTLAVDYPVKVSCYVSDDG  
Db 274 LSTLALDYPVKVSCYVSDDG  
Qy 437 KIDYLDKQVPSFVKDRAMK  
Db 334 KIDYLDKQVPSFVKERRAMK  
Qy 497 DHPGMIOVFLGHSGGLDTEG  
Db 394 DHPGMIOVFLGYSRGARDIEG  
Qy 557 YMLNLCDDHYNNKSKAVREAM  
Db 454 FILNLCDDHYNNKSKAVREAM  
Qy 617 INLRGLDGIQGPVYVGTGCVF  
Db 514 VNMKGLDGIQGPVYVGTGCVF  
Qy 677 KKSNKHVDSSVFPVNLIEDIE  
Db 574 LYDRAKREELDAALFNREID

ytic chain celA1 - upland cotton  
n 16-Jul-1999 #text\_change 21-Jul-2000  
ost, W.E.; Delmer, D.P.; Stalker, D.M.  
37-12842, 1996  
gs of the bacterial celA genes encoding the cata  
7296; PMID:8901635  
m GB/EMBL/DDBJ  
g1706955; PIDN:AAB37766.1; PID:g1706956  
U-2; fiber  
sis of cellulose  
syltransferase  
re 3477; DB 2; Length 974;  
ad. No. 1.3e-246;  
Mismatches 176; Indels 106; Gaps 15;  
VCRPCYEHKEKGTQACLOCKTKYKRRHGSPAI 77  
ICKSCFEYDLKEGRKACLR-----GSPY- 58  
KQKTADMRSPRMNTGGSGNVG-HPKYDSGIGL 136  
STMAA-----HLNKSQDVGIIHARHS---SV 96  
GASPDHMHMSPTGNISRRAPPYVNHSPNSREF 196  
-----EM 103  
TNTGTSIAPSEGRAADTDASTEYNMEDALLNDE 256  
ATT-----KVERAEIPEQOQMEDKPADPA 154  
LVLSIFLHYLTNPVRNAYPLWLVSICEIWF 316  
LIIIGLFEHYRVTNPVDSAFGLWLTSTVCEIWF 213  
VDREGEPSQLAAVDIFVSTVDPLEKPPITANTY 376  
VEREGEPDELAADVFFVSTVDPLEKPPITANTV 273  
PDALAEYSEFARKWVPFKYKDYIEPRAPERFECQ 436  
FESLVETADFAARKWVPFKKFSIEPRAPFYSQ 333  
FKIRINALVSKALKVPEBGWIMQDGTWPGNNT 496  
FKIRINALVAKAQKTDEGWTMQDGTWPGNNT 393  
VYVSRKRPFGQHHKAGAMNALRVSAVLNTNGQ 556  
VYVSRKRPFGQHHKAGAMNALRVSAVLNTAP 453  
PNLGPQVCYVOPQRFEDGIDRNDRYANRNTVF 616  
DQVGRDVCYVOPQRFEDGIDRNDRYANRNTVF 513  
YGEPPYKAKPGFLASLCGGKKKASKSKRSDD 676  
YGYGPPSPSPKSSSSSSSCSCCCKPKKPKDPSE 573  
AGFDD-EKSVLMSONSLEKRFQGSAAVASTLME 735  
-NYDEYKSLMISQTSFEKTFGLSVFTIESTME 627



QY 736 YGVPQSPSPSLLKEAIIHVISCYEDKSEWGTGIIWYGSVTEIDILTGKMHARGRSV 795  
DB 628 NGGVAESANPTSLKEAIIHVISCYEDKSEWGTGIIWYGSVTEIDILTGKMHARGRSV 687  
QY 796 YCMRPAFAKGSAPINLSDRLNQVLRWALGSVEILFRRHCPMLVGY-GGRLEKLEFAYI 854  
DB 688 YCMPLRPAKGSAPINLSDRLHQVLRWALGSVEILFRRHCPMLVGYGGGRLKWLQRLAYI 747  
QY 855 NTIYPLSLPLLYCIIIPALCLITGKFIPEISNLASIFIALFISIFATGILEMRWSG 914  
DB 748 NTIVYPTSLPIAYCSIPAICLTGKFIPTLSNLASVILFGLFISIIYAVTAVLEMRWSG 807  
QY 915 VGIDSWRNEQFWIIGGISAHLFAVFCGLKVLGADITNFTVTSKANDSGDFAELMYMK 974  
DB 808 VSIEDLRNEQFWIIGGISAHLFAVFCGLKVLGADITNFTVTSKANDSGDFAELMYMK 866  
QY 975 WTLIIPPTLIIINMGVWAGTSYAINSGVQSWGPGFLGKLFAPFVWIVHLYPFLKGLMG 1034  
DB 867 WTLIIPPTLIIINMGVWAGTSYAINSGVQSWGPGFLGKLFAPFVWIVHLYPFLKGLMG 926  
QY 1035 QNRPTTIVMAVLASIFSLWVRVDPFTTLAGPNI-QTC-GINC 1080  
DB 927 QNRPTTIVMAVLASIFSLWVRVDPFTTLAGPNI-QTC-GINC 974

## RESULT 9

T04870  
cellulose synthase (EC 2.4.1.-) catalytic chain F28A21.190 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-Jun-1999  
C:Accession: T04870  
R:Devan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer,  
submitted to the Protein Sequence Database, February 1999  
A:Reference number: 215387  
A:Accession: T04870  
A:Molecule type: DNA  
A:Residues: 1-958 <BEV>  
A:Cross-references: EMBL:AL035526  
A:Experimental source: cultivar Columbia; BAC clone F28A21  
C:Genetics:  
A:Map position: 4  
A:Introns: 60/1; 76/3; 94/1; 149/2; 240/2; 355/3; 401/3; 443/3; 514/3; 662/3; 780/3; 791/3  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 57.2%; Score 3302.5; DB 2; Length 958;  
Best Local Similarity 59.5%; Pred. No. 8.2e-234;  
Matches 638; Conservative 125; Mismatches 174; Indels 135; Gaps 17;

QY 18 VQICADGLGTTLDGVTACDVCRRPCYERHERKEGTQACLOCKTKYKRRHGPSPI 77  
DB 8 ICNTGCEITGVKSNGEFFVACHCECFPCACLEVEFEKGRRCICLRCGNPY----- 58  
QY 78 RGEEDDTDDGSDFNYPASTEDQKOKIADMRSMWMTGGSGNGVHPKYSDEIGLS 137  
DB 59 -----DENVFDDVE-----TKTSKQSID-----SGI--HARHIS---TVS 89  
QY 138 KYDSGEIPRGVPSVTNSQMSGEIPGASPDHMHMSPTGNISRRAPPFPYVNHSPNPREFS 197  
DB 90 TIDS-----ELN 96  
QY 198 GSTGNVANKERVDGWMKODKAIPWTNGTSTAPSEGRAANDIDASTEYNMEDALLND 257  
DB 97 DEYGNPIWNRVSWRDKKSKKKKDDPKATRAEQHEA---QIPQOHMEDTPPPNTES 153  
QY 258 --RQPLSRKVPYASSKINPYRMVIVLRVLVLSIFLHRLTNPNVRNAYPLWLLSVICBIWF 315  
DB 154 GATDVLVSVIPRTPKITSYRIVIMRLIILALFPNRYTHPVDYSAYGLWLTSVICBIWF 213  
QY 316 ALSWILDQPKFPFPNRETYIDRLALRDRECEPSQLAAVDIFVSTVDPLKEPPIVANT 375  
DB 214 AVSWLDQPKFPNRETYIDRLSARPEREGEQSLAAVDFFVSTVDPLKEPPIVANT 273

QY 376 VLSILAVDYPVDKYSCYVSDGASMLTFDALAETSEFARKWVPVKKYDIEPAPPEFC 435  
DB 274 VLSILALDYPVDKYSCYVSDGASMLTFDALAETSEFARKWVPVKKYDIEPAPPEFC 333  
QY 436 QKIDYLDKQVSPFVKORRAMKREYEEFKIRINALYSKALKVPPEGIMODGTPWPGNNT 495  
DB 334 LKIDYLDKQVSPFVKORRAMKREYEEFKIRINALYSKALKVPPEGIMODGTPWPGNNT 393  
QY 496 RDHFGMIQVFLGHSGGLDTFEGNELPLVYVSRKRCFOHKKAGAMNALVRVSAVLNMG 555  
DB 394 RDHFGMIQVFLGHSGGLDTFEGNELPLVYVSRKRCFOHKKAGAMNALVRVSAVLNMG 453  
QY 556 QYMLNLCDDHYINNSKAVREACFLMDPNLGPQVYVQFPQRFQREDGIDRNDRYANRNTVFF 615  
DB 454 PFILNLCDDHYINNSKAVREACFLMDPNLGPQVYVQFPQRFQREDGIDRNDRYANRNTVFF 513  
QY 616 DINLRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFL-----ASLCGKKKASKS 670  
DB 514 DVNMRGLDGIQGVYVGTGCVFNRTAIYGYEPPKAKKPGFL-----ASLCGKKKASKS 568  
QY 671 KRRSDDKKKSNKHYDSSVPFENLEDIEEGVEGAGFDD-EKSVLSMSQMSLEKREFGQSAFV 729  
DB 569 PQDPSEIYKDAKREELDAAIFNLGDL-----NYDEVDRSMLISQTSPEKTFGLSTVFI 622  
QY 730 ASTLMEYGGVQSPSTPESLLKEAIIHVISCYEDKSEWGTGIIWYGSVTEIDILTGKMH 789  
DB 623 ESTLMENGSGVPSYNSPTSLKEAIIHVISCYEDKSEWGTGIIWYGSVTEIDILTGKMH 682  
QY 790 RGRSVYCMKRRPAFAKGSAPINLSDRLNQVLRWALGSVEILFRRHCPMLVGY-YGGRLEK 848  
DB 683 RGRSVYCMKRRPAFAKGSAPINLSDRLHQVLRWALGSVEILFRRHCPMLVGY-YGGRLEK 742  
QY 849 ERFAVINTIYPLSLPLLYCIIIPALCLITGKFIPEISNLASIFIALFISIFATGIL 908  
DB 743 ORLAYINTIYPTSLPIAYCSIPAICLTGKFIPTLSNLASVILFGLFISIIYAVTAV 790  
QY 909 EMRWGSGVIDEWRNEQFWIIGGISAHLFAVFCGLKVLGADITNFTVTSKANDSGDFA 968  
DB 791 -----GVSIEDLRNEQFWIIGGISAHLFAVFCGLKVLGADITNFTVTSKANDSGDFA 844  
QY 969 ELYMKFWTTLIIPPTLIIINMGVWAGTSYAINSGVQSWGPGFLGKLFAPFVWIVHLYP 1028  
DB 845 ELYMKFWTTLIIPPTLIIINMGVWAGTSYAINSGVQSWGPGFLGKLFAPFVWIVHLYP 904  
QY 1029 LKGLMGRQNRPTTIVIVMAVLASIFSLWVRVDPFTTLAGPNIQTCGINC 1080  
DB 905 LKGLMGRQNRPTTIVIVMAVLASIFSLWVRVDPFTTLAGPNIQTCGINC 953

## RESULT 10

T10800  
cellulose synthase (EC 2.4.1.-) catalytic chain celsa2 - upland cotton (fragment)  
C:Species: Gossypium hirsutum (upland cotton)  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T10800  
R:Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.  
Proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996  
A:Title: Higher plants contain homologs of the bacterial celsa genes encoding the cat-  
A:Reference number: #17152; MUID:97057296; PMID:8901635  
A:Accession: T10800  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-685 <PEA>  
A:Cross-references: EMBL:U58284; NID:gl706957; PIDN:AAB37767.1; PID:gl706958  
A:Experimental source: strain Acala Su-2; fiber  
C:Genetics:  
A:Gene: celsa2  
C:Function:  
A:Description: involved in synthesis of cellulose  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 48.7%; Score 2816; DB 2; Length 685;  
Best Local Similarity 74.0%; Pred. No. 2.5e-198;



Matches 509; Conservative 78; Mismatches 77; Indels 24; Gaps 6;

```
QY 414 RKWPFVKKYDIEPRAPFEYFCQIDYLDKQVPSFVKDRAMKREYEEFKIRINALVSK 473
D 1 RWWPFCKKHNVPRAPFEYNEKIDYLDKQVPSFVKERRAMKREYEEFKIRINALVAK 60
QY 474 ALKVPPEGWIMODGTPWGNTRHPGMIOVFLGSHSGGLDTFEGNELPLVTVVSEKRPGF 533
D 61 AOKKPEEGWIMODGTPWGNTRHPGMIOVFLGSHSGGLDTFEGNELPLVTVVSEKRPGF 120
QY 534 QHKKKAGAMNALVRYSALVTNGQYMLNLDCHYINNSKAVREAMCFLMDPNLGPQVCVQV 593
D 121 QHKKKAGAMNALVRYSALVTNAPFLNLDCHYINNSKAVREAMCFLMDPNLGPQVCVQV 180
QY 594 PQRFDGIDRNDRYANRNVTFDNLRLGDIQGPVYVGTGCVNRTAIYIYEPPIKAK 653
D 181 PQRFDGIDRNDRYANRNVTFDNLRLGDIQGPVYVGTGCVNRTAIYIYEPPIKAK 240
QY 654 -----PGFLASLCGGKKKASKSKRS-----SDKKK-----SNKHVDSVPVFN 693
D 241 PKMTCDCHPSWCCCCCGSRKSKKKGKGLLGLLGGKMMKKNYVKKGSAFVFDL 300
QY 694 EDIEBVEGAGFDD-EKSVLMSOMSEKRFQCSAAFVASTLMEXGGVPQSSPTESLLKEA 752
D 301 EBIIEGLE--GYEELEKSTLMSQKNFKRFQSPVFIASITLMENGLPEGTNSTSLIKEA 358
QY 753 THVISCYGEDKSEWTEIGWYGVSTEDILTGFKWHARGWSVYCMKRPAPKGSAPINL 812
D 359 THVISCYGEETKEMWKEIGWYGVSTEDILTGFKWHARGWSVYCMKRPAPKGSAPINL 418
QY 813 SDRNLQVLRWALGSVEILFSRHCPLMWYGGYGRKLFERFAYINTIYPLTSLPLLVYCIL 872
D 419 SDRNLQVLRWALGSVEILFSRHCPLMWYGGYGRKLFERFAYINTIYPLTSLPLLVYCIL 478
QY 873 PAICLLTGKFMPEISNLASTWFTALFLSIFATGILEMRWSGVGIDEMRNEQFWIGGI 932
D 479 PAICLLTGKFIPTISNTSWFALFLSIIATGYLELRWSGVSIQDWMRNEQFWIGGV 538
QY 933 SAHLFAVQGLLKVLAGIDTNTFTVTSKANDEGDFAEILYMFKWTLLIPPTIILINNVG 992
D 539 SAHLFAVQGLLKVLAGIDTNTFTVTSKANDEGDFAEILYMFKWTLLIPPTIILINNVG 597
QY 993 VVAGTSYAINSGYOSGWLPGKLFPAFWIVHLYPFLKGLMGRNRTPTIIVWAVLIAS 1052
D 598 VVAGTSYAINSGYOSGWLPGKLFPAFWIVHLYPFLKGLMGRNRTPTIIVWAVLIAS 657
QY 1053 IFSLLWVRVDPFTFLAGNPOTGCGINC 1080
D 658 IFSLLWVRVDPFTFLAGNPOTGCGINC 685
```

RESULT 11

D86157  
hypoetical protein F22D16.26 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 sequence revision 02-Mar-2001 #text\_change 31-Dec-2001

R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86157

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1181 &lt;STO&gt;

A:Cross-references: GB:AE005172; NID:g6056428; PIDN:AAF02892.1; GSPDB:GN00141

C:Genetics:  
A:Map position: 1Query Match 40.1%;  
Best Local Similarity 42.9%;  
Matches 488; Conservative 1

```
QY 15 AGDVCOI--CADGLGTITLDGDV 40.1%;
D 158 SGQICWLKGGCKEY-----V 42.9%;
QY 73 GSPAIRCEEGDDTDADDGSDFN 488;
D 210 DDP-----ETEEED----- 488;
QY 133 EIGLSKYDSGEIPRGVYPSVTN 40.1%;
D 228 QMGESKLDK---RLSVVVKFKA 42.9%;
QY 193 SREPSGSITGNVANKERVGDGKK 488;
D 269 -----GNAYWPK--DGYG 488;
QY 253 LNDETROPRLSRKVPDIASSKINE 40.1%;
D 297 --ERSKRPLTRKVSVAIIIS 42.9%;
QY 313 IWFALSILDOFPKWFPIINRE 40.1%;
D 355 LWFALSILDOFPKWFPIINRE 42.9%;
QY 368 PPIVTANTVLSILAVDYPVK 40.1%;
D 415 PPIVTANTVLSILAVDYPVK 42.9%;
QY 428 RAPEYFCQIDYLDKQVPS 40.1%;
D 475 RNPEYFCQIRNFKVRLD 42.9%;
QY 476 K-----V 40.1%;
D 535 RAKKKQEMMMGNPNQETVIV 42.9%;
QY 509 SGGLDTEGNE----- 40.1%;
D 594 PNAEPVYGAEDAENLDTTD 42.9%;
QY 554 NGQYMLNLDCHYINNSKAVR 40.1%;
D 654 NGPFILNLDCHYIYNSMALR 42.9%;
QY 614 FFDINLRGLDGIQGPVYVGTG 40.1%;
D 713 FFDVSMRALDGLQGMVGTG 42.9%;
QY 674 SSDKKKNKHVDSSVPV---F 40.1%;
D 768 PKAMKKKDDVE--SLPINGEY 42.9%;
QY 731 S-TLMEXYGG----- 40.1%;
D 814 SIPVAEYQGRLIQDLOGKGN 42.9%;
QY 770 IGWYIGSVTEIDITGFKMHAR 40.1%;
D 874 VGVYIGSVTEIDITGFKMHAR 42.9%;
QY 830 LFSRHCPLMWYGGYGRKLFER 40.1%;
D 934 FFSRNNAIF--ATPRMFLOR 42.9%;
QY 890 LASTWFTALFLSIFATGILEM 40.1%;
D 934 LASTWFTALFLSIFATGILEM 42.9%;
```

re 2315; DB 2; Length 1181;

d. No. 3.4e-161;

Mismatches 256; Indels 214; Gaps 31;

VCRPPVCRPCYHERKEGTQACLOCKTKYKRRH 72

-CGFRICRDCYFCITGGGCGCKEYPYRDIN 209

STEDQKOKIADRMRSWRMNTGSGNVGPKYDSG 132

EEDAKPLP----- 227

SEIPGASPDHMHMSPGTGNISRRAPPPVYVNHSPN 192

SDF-----DHPRWL-----FETKGYGY 268

SAIPMTNCTSTAPSEGRAANTDIDASTEYNMEDAL 252

SG-----NGYETPPEFG----- 296

IVLRVLVLSIFLHYRLTNPNVRNAYPLWLLSVCE 312

FALRLVGLGLFLTWVRHPNREAMWLMCMSTTCE 354

LALRYDREG--EP--SOLAADVIFSVTVDPKE 367

LKERFESPNLRNPKRSDLPGLDIFVSVTADPEKE 414

SDDGASMLTFDALAETSEFARKWVPFVKKYDIEP 427

SDDGALLTFEALAQTFASFASTWVPFCRKHIEP 474

SAMKREYEEFKIRINAL-----VSKAL 475

SRVKREYDEFRVINSLSPEAIRRSRSDAYNVHEEL 534

TMDGTMPG-----NNTRDHPGMIQVFLGH 508

MSDGSHPGWTGTSGETDNGSDHAGIIQAMLAP 593

PRLVYVSEKRPFGQHHKKAGAMNALVRVSAVL 553

PMLVYVSEKRPGYDHNKAGAMNALVRTSAIMS 653

LMDPNLGPQVCYVQFPQRFQDIDRNDRYANNTV 613

MLDRG-GDRIYVQFPQRFQDIDRNDRYANNTV 712

TAIYGEPPPIKAKKPGFLASLCGGKKKASKSKR 673

TALYGFSPPRATEHHGWL-----GRRKVKISLR 767

TEEGVEGAGFDDKSVLMSOMSELEKRFQCSAAFVA 730

DDG-----DIESLL-----LPRKFGNSNFEVA 813

---VQSSPTSPBSLLKEATHVISCYGEYDKSEWTE 769

SLAVPREPLDAAVTAELSVLSCEYEDKTEWGR 873

YCMKPRPAFKGAPINLSDRNLQVLRWALGSVEI 829

YCVTKRDAFRGTAPINLTDRLHQLWRWATGSVEI 933

VTIYPLSLPLLVYCIIPALICLLTGKFMPIISN 889

VGMYPFTSLFVYCIIPALISLFGQFIVQSLDI 991

GIDEMRNEQFWIGGISAHULFAVFGGLKVLKLAG 949

Db 992 TELIYLLSITLTLCLMSLLEIKWSGITLHEWRNEQFWVIGTSAHPAAVLQGLLKVIAG 1051  
QY 950 IDTNETVTSKAN-DEEGD--FAELYMFKWTTLLIPPTTILLINNVGVVAGTSYAINSGYQ 1006  
Db 1052 VDISFTLTKSSAPEDGDEADLXVVKWSELMVPLTIMVMNIAIAGLARTLYSFPF 1111  
QY 1007 SGMPLFGKLFAPFWIVHLYPFLKGLMGQRNRTPIVIVWAVLASIFSLLLWVRVDP 1063  
Db 1112 QMSKLVGGVFFSEFWLCHLYPFAKGLMGRGRVPTIVFWSGLLSIIYVSLLLWYINP 1168  
RESULT 12  
T05646  
hypothetical protein F20D10.310 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999  
C:Accession: T05646  
R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.H.  
submitted to the Protein Sequence Database, February 1999  
A:Reference number: 215420  
A:Accession: T05646  
A:Molecule type: DNA  
A:Residues: 1-1111 <BEV>  
A:Cross-references: EMBL:AL035538  
A:Experimental source: cultivar Columbia; BAC clone F20D10  
C:Genetics:  
A:Map position: 4  
A:Introns: 139/2: 675/3  
A:Note: F20D10.310  
Query Match 39.1%; Score 2257; DB 2; Length 1111;  
Best Local Similarity 41.8%; Pred. No. 5.5e-157;  
Matches 486; Conservative 159; Mismatches 273; Indels 244; Gaps 32;  
QY 4 DADALKSRHGA-GDVQCICADGLTTLTLDGVF-----TACDYCRFPVCRPCYEHR 54  
Db 99 DSDVTHPQAGAKGSSCAMPA-----CDGNVMDKRGKDVMPCE-CRFKICRDCFMDAQ 151  
QY 55 KEGTOACLOQCKTKYKRHRGSPAIRGEEDDTDADDGSDFNYPASGTEDQKQIADRMRSW 114  
Db 152 KE-TGLCPCKEQYK-----IGLDLDDTDPDYSSGALPLPAPG-KDQR----- 191  
QY 115 RMNTGGNGVHPKYDSGEIGLSKYDSGEIPRGYYPVTNSQMSGEIPGASPDHMHMSPT 174  
Db 192 ----GNNNM-----SMKRNQGEF-----DHNRL-- 214  
QY 175 GNISRAPPPYVHNPSRSESGIGNVANKERYDGKMKQDKCAIPWNTGTSTAPSEG 234  
Db 215 --FETQGTGY-----GNAYW-----PQDE 232  
QY 235 RAATDIDASTEYNMEDALLNDTROPLSRKVPFIASSKINPVMIVLRLVLSIFLHVRL 294  
Db 233 MYGDDMDEGMRCGMVETA--DKPWPLSRRIPIPAALISPYRLLIVIRFVLCFLTWRI 290  
QY 295 TNPVRNAYPLWLLSVCEIWFALSILDOFPKWFPIINRETYLDRALRYDREGEP----- 349  
Db 291 RNPNEDAIWLWLSIICELWFGEFWILQIPKLCPIINRSTDLVLRDKDFPMSPSNPTGR 350  
QY 350 SOLAAVDVFSVIVDPLKPPPIVANTVLSILAVDPVKVSCYVSDDGASMLTFDALAET 409  
Db 351 SLDPLGIDLFVSADPEKEPPLVTANTILSLAVDPYVEKVSICYLSDDDGGLISFEAMAEA 410  
QY 410 SEFARKWVPVKKYDIEPRAPEFYFCOKIDYLDKDVQVSEFKVDRAMKREVEEFKIRINA 469  
Db 411 ASFADLWPFCKRHNIERNPDSYFSLKIDPTKNKSRIDFVKDRKIRREYDEFKVRING 470  
QY 470 LVS-----KAL-----KVPPEGWIMQDGTWPFCN-----NTR- 496  
Db 471 LPDSIRRRSDAFNAREEMKALKQMRSGSDPTEPVKVPKATW-MADGTHWGTWAASTRE 529  
QY 497 ----DHPGMIOVFL-----GHSGG-----LDTEGNELPRLVYVSRKRPQGHKK 538  
Db 530 HSKGDHAGILQWMLKPPSPSDPLIGNSDKVIDFSOTD-TRLPMFYVYSRKRPGVDHNK 588

QY 539 AGAMNALVRYSAVLTNGOYMLNLDCHYINNSKAVREAMCFMLDPNLPQVCYVQPPORF 598  
Db 589 AGAMNALVRASAILNSGFFILNLDCHYIYNCKAVRGCMFMDRG--GEDICITQPPORF 647  
QY 599 DGDNRDRYANRNTVFEDINLRGLDGIQGPVYGTGCVFNRTAIYGYEPPKAKKPGFLA 658  
Db 648 EGIDPSRYANNTVFFDGNMRALDVGQPVYGTGTMFRREALYGFDDP----- 697  
QY 659 SLGGGKKKASKRRSSDKKSNKHVDSSVPVFNLEIDIEGVEGAGFDDDEKSVLMQMSL 718  
Db 698 -----NPKDLLEKK-----ESETALATSDFDPLDV---TQL 727  
QY 719 EKRFQSOAAFWAS-TLMEYGG-----VPOSSPTESLLKEAHRVISC 758  
Db 728 PKRFGNSTLLAESIPIAEFOGRPLADHPAVKYGRPPGALRVPRDPLDATTVAESVVIS 787  
QY 759 GYEDKSEWGTGIEWIGYSVTEDILTGFKMARGWRSYCMKPAFKGSAPIINLSRLNQ 818  
Db 788 WYEDKTEWGDVGVWIGYSVTEDVVTGYRMHNRGWSYVYITKRDSPRGSAPINLTORLHQ 847  
QY 819 VLKHALGSVELLSRHCPLWYGGYGRLLKFLERFAYINTTIYPLTSLPLLYCYLPAICLL 878  
Db 848 VLKMATGSVEIFFSRNAIL--ASKRLKFLORLAYLVNGYIPFTSLFLILYCFLPAPSLF 905  
QY 879 TGRFIMPEISNLASINFIATLSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISAHLEA 938  
Db 906 SGQFIVPTLSIFLVYLLMITICLIGLAVLEKWSGIGLEWRNEQWMLISGTSULYA 965  
QY 939 VFQGLLKVLAGIDINFTVTSKA--NDEEGDFAELYMFKWTTLLIPPTTILILINMVGVAG 996  
Db 966 VVQGLVKVIAIEISFTLTTKGGDDNEDIYADLYIVKWSLMPPIVIAWNIATVWA 1025  
QY 997 TSYAINSGYSWGLFCKLFFAFWVIVHLYPFLKGLMGQRNRTPIVIVWAVLASIFSL 1056  
Db 1026 FIRTIOAVPQWSKLGAGFFSEFWLHLYPFAKGLMGRGRKTPITIVFWAGLIAITISL 1085  
QY 1057 LWRVVDPTTRLAGPNITQTCGI 1078  
Db 1086 LWTAINPNT---GPAAAAEV 1103  
RESULT 13  
T51546  
cellulose synthase catalytic subunit-like protein - Arabidopsis thaliana  
N:Alternate names: protein F2K13\_60  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 18-Aug-2000  
C:Accession: T51546  
R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.;  
submitted to the Protein Sequence Database, August 2000  
A:Reference number: 225394  
A:Accession: T51546  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1145 <SAT>  
A:Cross-references: EMBL:AL391141  
A:Experimental source: cultivar Columbia; BAC clone F2K13  
C:Genetics:  
A:Map position: 5  
A:Introns: 297/2: 566/3  
A:Note: F2K13\_60  
Query Match 38.4%; Score 2220.5; DB 2; Length 1145;  
Best Local Similarity 42.4%; Pred. No. 2.8e-154;  
Matches 479; Conservative 166; Mismatches 280; Indels 205; Gaps 27;  
QY 16 GDVCOI--CADGLTTLDDGVFTACDVCRPVCRPCYEHKKECTQACLOCKTKYKRHRG 73  
Db 125 GSSCAIPGCAKVMDSRGGQDLIPCE--CDFKICRDCFDVAVKYGCGGICPCKEPPYK---- 179  
QY 74 SPAIRGEGDDTDADDGSDFNYPASGTEDQKQIADRMRSWRMNTGSGVNGVHPKYDSGE 133



Db 651 LNDPDP-----LGLPKKFGNSTMTDTPVAEYOGRLADHMSVKNRPPGALLPRP 703  
QY 750 -----KEAHIVISGDEYDKSWGTEIGWYGSVTEIDLTFKMHARGWSRYVCMKRP 802  
Db 704 PLDAPTVAEIAVISCWYEDNTGWRIGWYGSVTEVDVTVGYRMHNRGWSRYVCTKRD 763  
QY 803 AFKGSAPINSLDRNLGVRWALGSVELTFSRCHPCPLAWYGGYGRLLKFLERFAYINTTYPLT 862  
Db 764 AFRGTAPINTDRLHQVLRWATGSVEIFFSKNNMF--ATRRUKFLQVAYLNVGIVPFT 821  
QY 863 SLPLVYCILPAICLTGKFMPEISINLASINFIALFLSIFAT-----GILEMRWSGVGD 918  
Db 822 SIFLVYCFPALCLFSGKFIQVSL-----DIHFLSYLLCITVTTLISLLEVKWSGIGLE 877  
QY 919 EWRNEQFWWIGISAHFLFAVFOGLLKVLAGIDTNTVTYSKANDERG--FAELYMFKWT 976  
Db 978 EWRNEQFWWIGISAHFLFAVFOGLLKVLAGIDTNTVTYSKANDERG--FAELYMFKWT 937  
QY 977 TLLIPPTTILINNMGVWAGTSVAINSGYSGWGLFGKLFFAFWVIVHLYPFLKGLMGRO 1036  
Db 938 GLFIMPFTIILNVAIVGASRTIYVIPQWCKLGGIFFLVWVTHMYPFAKGLMGRR 997  
QY 1037 NRTPTIVVWAVLLASIFSLWVRDP 1063  
Db 998 GKVPTIVVWGLSVITVSLWITISP 1024

RESULT 15  
C86446  
probable cellulose synthase catalytic subunit [imported] - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C/Accession: C86446  
R/Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Dewar, T.H.; Dewar, K.;  
Nature 408, 815-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: C86446  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-979 <STO>  
A/Cross-references: GB:AE005172; NID:g10801364; PIDN:AAG23436.1; GSPDB:GN00141  
C/Genetics:  
A/Map position: 1

Query Match 37.4%; Score 2163.5; DB 2; Length 979;  
Best Local Similarity 44.3%; Pred. No. 3.3e-150;  
Matches 450; Conservative 173; Mismatches 245; Indels 147; Gaps 24;

QY 142 GEIPRGVPSVTSNMSGETPGASPDHMHMSPTGNTSRAPPYVNHSPNPSREFSGSIG 201  
Db 4 GESPLRH-PRISHVNSGSGFGSSSYNKY-----LVQIPPTP--DNNGPA-----SUS 50  
QY 202 NVAWKERVGKMKQDKGAIPMTNGTSIAPSEGR-----AATDIDASTEYNMEDALLNDET 257  
Db 51 IVL-----LEIDSQESVPSVGDIVSGSGCKDNEPOLTVRINVGEEDDDTLSSKI 103  
QY 258 RQPLSRKVPITASKINPYRVVLRVLVLSIFLHYRLTNPNRNAYPLWLLSVCEIWFAL 317  
Db 104 SYSLTRVVKISPIIALYRILIVVRVSLALFLWRIRPNPNKALWLWLLSVCEIWLWAF 163  
QY 318 SWLLDQPKWPEPNRETYDLRALRY-----DREGESQAAVDIEFVSVDPLKEPIVTV 372  
Db 164 SWLLDQPKWPEPNRETYDLRALRY-----DREGESQAAVDIEFVSVDPLKEPIVTV 223  
QY 373 ANTVLISILAVDYPVDKVCYSVSDGASMLTFOALAESEFARKWVPVFKKYDIEPRAPEF 432

Db 224 ANTILISLUSVDYPEKLSVYISDDGSLVTEFAETAASAFKIWWPFCRKHKIEPRNPES 283  
QY 433 YFCQKIDYKQVQSFYKDRAMKREYEEFKIRINALV-----SK----- 473  
Db 284 YFGLKRDYKDKVRHDFVRRERYVKRAYDEFKVRNALPHSIRRSADFNSKEEIKALEK 343  
QY 474 -----ALKVPEEGIMQDGTWPQG-----NNTRDHPGMIQVFL- 506  
Db 344 WKHKWKVVEEDQIKRPRPALVAPKATW-MSDGTWHPGTWAYSGPHHSGDHSYIQVLLD 402  
QY 507 -----GHSGLLDEGNE--LPRLVVVSREKRRGFOHKKKAGAMNALVRSVAVLT 553  
Db 403 PGDEPVECKGEGALDLEGVDLRPLMLVVSREKRRGFOHKKKAGAMNALVRSVAVLT 462  
QY 554 NGQVMLNDCDHYINNKAVERAMCFLMDPNLGPQVCYVQPPQRFQDIDRNDRYANRTV 613  
Db 463 NGPFTLNLDCHYVYNSRAFRDGCIFMMDHD-GDRVSYVQPPQRFQDIDRNDRYANRTV 521  
QY 614 PFDINRLGDLGTOGPPVYVGTGCVENRTAIYGEPP---IKAKKPGFLASLGGGKKASKS 670  
Db 522 PFDINRLGDLGTOGPPVYVGTGCVENRTAIYGEPP---IKAKKPGFLASLGGGKKASKS 574  
QY 671 KKRSSDKKKSNKHYDSSVPFVFNLEDERGVEGAGDFDDEKSVLMSQMSLEKRFQSOAAFA 730  
Db 575 KKRSPATVAS-----EPEYTTDEEDRFDIGL-IRKQFGSSMLVN 613  
QY 731 ST-LMEYGGVQSSST-----PESL-----LKEAHIVISGDEYDKSWGTEI 770  
Db 614 SVKVAEFGRRPLATVHSSRLGRPPGSLTGSRKPLDFATVNEAVNVISWYEDKTEWGFNV 673  
QY 771 GWYGSVTEIDLTFKMHARGWSRYVCMKRPFAKGSAPINSLDRNLGVRWALGSVELL 830  
Db 674 GWYGSVTEIDLTFKMHARGWSRYVCMKRPFAKGSAPINSLDRNLGVRWALGSVELL 733  
QY 831 FSRHCPLWYGGGRLLKFLERFAYINTTYPLTSLPLVYCILPAICLTGKFMPEISINL 890  
Db 734 FSRNNAIF--AGPKLKLQRIAYLNVGIVPFTSIFILTYCFLPPLSLFSGHFVETLGS 791  
QY 891 ASINFIALFLSIFATGILEMRWSGVGDIEWRNEQFWWIGISAHFLFAVFOGLLKVLAGI 950  
Db 792 FLIYLLIITLSLCLGLAVLEVKWSGISLEWRNEQFWWIGISAHFLFAVFOGLLKVLAGI 851  
QY 951 DTNFTVTSKA---NDEEGDFAELYMFKWTLLIPPTTILINNMGVWAGTSVAINSGYQ 1006  
Db 852 EISFTLTSKSTGSGDDEDEDFADLYLFWKATMLPPLIILINIVAILFAVCRVTFVSANP 911  
QY 1007 SWGPFLGKLFFAFWVIVHLYPFLKGLMGRQNRTPPTIVVWAVLLASIFSLWVRV 1061  
Db 912 QWSNLLGGTFFASWVLLHMYPFKGLMGRGKTPVTVVWVWSGLIAICLSLIYITI 966

Search completed: February 19, 2003, 16:42:41  
Job time : 59 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: February 19, 2003, 16:34:11 ; Search time 34 Seconds  
(without alignments)  
1317.483 Million cell updates/sec

Title: US-09-900-237-30

Perfect score: 5778

Sequence: 1 MDGADALKSRHGAGDVCO.....VDFFTRLAGPNIQTCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	264	4.6	874	1	BCSA_SALTI
2	264	4.6	874	1	BCSA_SALTY
3	259.5	4.5	1596	1	ACS2_ACEXY
4	258	4.5	872	1	BCSA_ECO57
5	257	4.4	872	1	BCSA_ECOLI
6	248	4.3	739	1	BCSA_PSEFL
7	238.5	4.1	754	1	BCAL_ACEXY
8	238.5	4.1	1518	1	BCAM_ACEXY
9	237.5	4.1	1518	1	BCA5_ACEXY
10	230	4.0	729	1	BCSA_XANAC
11	220	3.6	1550	1	ACS1_ACEXY
12	207	3.5	745	1	BCA3_ACEXY
13	202.5	3.5	745	1	BCA3_ACEXY
14	119	2.1	1433	1	SUBF_BACSU
15	116.5	2.0	441	1	YCDQ_ECOLI
16	112.5	1.9	322	1	NUIM_STRPU
17	112	1.9	1029	1	RIP3_RAT
18	111	1.9	414	1	YOAB_BACSU
19	110.5	1.9	1154	1	KDGD_MESAU
20	109.5	1.9	1024	1	RIP3_MOUSE
21	108.5	1.9	1262	1	TPP2_MOUSE
22	108	1.9	1167	1	CLAA_BACTU
23	108	1.9	1894	1	ME21_SCHPO
24	107.5	1.9	775	1	ICP0_HSV11
25	107.5	1.9	920	1	DMF1_SCHPO
26	107.5	1.9	1671	1	DPOL_PYRKO
27	107.5	1.9	3033	1	POLG_HCVJ6
28	107	1.9	1249	1	TPP2_RAT
29	106	1.8	357	1	G6PT_CANFA
30	105.5	1.8	662	1	CYOB_BUCAL
31	105.5	1.8	886	1	YKRL_CABEL
32	105	1.8	357	1	G6PT_HUMAN
33	104.5	1.8	541	1	COX1_PODAN

34 104.5 1.8 694 1 RP3A\_HUMAN Q9y2j0 homo sapien  
35 104.5 1.8 992 1 UVRA\_MICLU P13567 micrococcus  
36 104 1.8 374 1 MTLB\_BACHD Q9k681 bacillus ha  
37 104 1.8 517 1 COX1\_PARLI P12700 paracentrot  
38 104 1.8 1249 1 TPP2\_HUMAN P20439 drosophila  
39 103.5 1.8 530 1 CG2B\_DROME P24010 bacillus su  
40 103.5 1.8 622 1 COX1\_BACSU P17570 lycopersico  
41 103.5 1.8 911 1 NIA\_LYCES Q04723 lactococcus  
42 102.5 1.8 435 1 PEPC\_LACLC Q9y2g3 homo sapien  
43 102.5 1.8 672 1 A11B\_HUMAN Q9wvc7 rattus norv  
44 102.5 1.8 2314 1 AKA6\_RAT P33334 saccharomyc  
45 102.5 1.8 2413 1 PR08\_YEAST

## ALIGNMENTS

RESULT 1  
BCSA\_SALTI  
ID BCSA\_SALTI STANDARD: PRT: 874 AA.  
AC Q8z291;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).  
GN BCSA OR STY4181  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Salmonella.  
OX NCBI\_TaxID=601;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CT18;  
RX MEDLINE=21534947; PubMed=11677608;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
RA Whitehead S., Barrall B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
CC -!- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
uridine 5'-diphosphate glucose to cellulose, which is produced as  
an extracellular component for mechanical and chemical protection  
at the onset of the stationary phase, when the cells exhibit  
multicellular behavior (rod morphology). co-expression of  
cellulose and thin aggregative fimbriae leads to a hydrophobic  
network with tightly packed cells embedded in a highly inert  
matrix (By similarity).  
CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1,4-beta-D-glucosyl)](N) = UDP  
+ [(1,4-beta-D-glucosyl)](N+1).  
CC -!- COFACTOR: Magnesium (By similarity).  
CC -!- ENZYME REGULATION: Activated by bis-(3'-5') cyclic diguanylic acid  
(C-di-GMP) (By similarity).  
CC -!- PATHWAY: Bacterial cellulose biosynthesis.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
(Potential).  
CC -!- DOMAIN: There are two conserved domains in the globular part of  
the protein: the N-terminal domain (domain A) contains  
the conserved DxD motif and is possibly involved in catalysis and  
substrate binding. The C-terminal domain (domain B) contains the  
QXXRW motif and is present only in processive glycosyl  
transferases. It could be involved in the processivity function of  
the enzyme, possibly required for holding the growing glycan chain  
in the active site.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -





CC -!- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (potential).  
 CC -!- DOMAIN: There are two conserved domains in the globular part of  
 CC the protein: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC OXRRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AJ315770; CAC44015.1; -  
 CC DR EMBL: AE008667; AAL22479.1; -  
 CC DR EMBL: AJ315148; CAC86199.1; -  
 CC DR StyGene; SG2222; bcsA.  
 CC InterPro; IPR005150; Cellulose\_synt.  
 CC InterPro; IPR001173; Glycos\_transf\_2.  
 CC Pfam; PF03552; Cellulose\_synt; 1.  
 CC Pfam; PF00535; Glycos\_transf\_2; 1.  
 CC KW Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 CC Transmembrane; Inner membrane; Complete proteome.  
 CC FT DOMAIN 271 364 CATALYTIC SUBDOMAIN A.  
 CC FT DOMAIN 441 501 CATALYTIC SUBDOMAIN B.  
 CC FT TRANSMEM 30 50 POTENTIAL.  
 CC FT TRANSMEM 151 171 POTENTIAL.  
 CC FT TRANSMEM 173 193 POTENTIAL.  
 CC FT TRANSMEM 230 250 POTENTIAL.  
 CC FT TRANSMEM 525 545 POTENTIAL.  
 CC FT TRANSMEM 547 567 POTENTIAL.  
 CC FT TRANSMEM 592 612 POTENTIAL.  
 CC FT TRANSMEM 634 654 POTENTIAL.  
 CC FT TRANSMEM 668 688 POTENTIAL.  
 CC FT TRANSMEM 833 853 POTENTIAL.  
 CC FT ACT\_SITE 313 313 POTENTIAL.  
 CC FT ACT\_SITE 457 457 POTENTIAL.  
 CC FT SITE 360 360 SUBSTRATE BINDING (POTENTIAL).  
 CC FT SITE 362 362 SUBSTRATE BINDING (POTENTIAL).  
 CC SQ SEQUENCE 874 AA; 100043 MW; 4C9421B58606310A CRC64;

Query Match 4.68; Score 264; DB 1; Length 874;

Best Local Similarity 19.58; Pred. No. 1.2e-10;

Matches 159; Conservative 91; Mismatches 220; Indels 346; Gaps 30;

QY 261 LSRKVPFIASSINPRVIVLRLVLSITFLHYRLT-----NPVRNAYPLWLLSVICEIW 314  
 Db 189 LVRRNPGFSAI-----MLVLSLVSCRYINWRYTSTLNDPDPVSLVCGILLIFAETYAW 244  
 QY 315 FALSILDOFKPWFPPINRKYDRLALRYDRGEPGSAADIVFSTVDPPLKEPIIVTAN 374  
 Db 245 IVL--VLGYFQVWPLNRQ-----VPLPKEMSQWPTVDFIVPTYN---EDLNVRKN 291  
 QY 375 TVLSLITAVDPKVCVSDGASMLTFDALAETSEFAKWKVPFVKKYDIEPRAPEFFY 434  
 Db 292 TIYASLGIDWPKDLNIWILDDG----- 314  
 QY 435 COKIDYLDKRVQPSFVKDRAMKREYEEFKIRINALYSKALKVPEEGWIMQDGTWPNGNN 494  
 Db 315 -----GRE 317  
 QY 495 FPDHPGMIQVFLHSGGLDTGSELPRILVYVSREKRPQFQHHKKAGAMNALVRVSAVLTN 554  
 Db 318 SFRH-----FARHVG-----VHYIARTT-----HEHAKAGNINNALKHA----K 352

QY 555 GQVNLNDCDHY INNSKAVREAMCFELMDPNLGPQVCYVOPPORF-----DGIERN-----DRY 607  
 Db 353 GEFAIFDCDCHVTRSFLOMTMGWFLKE-----KQLAMMQTPHHFFSPDFERNLGRFKT 408  
 QY 608 ANRNVTFDDINLRGLDGIQGVYVGTGVNRTAIYGYEPIKAKKPGFLASLCGGKKKA 667  
 Db 409 PNEGTFLVGLVQDGNMDWDATFFCGSCAVIRR-----RP----- 442  
 QY 668 SKSKRRSSDKKSNKHVDSSVPVFNLEIEEGVEGAGFDDKSVLMSQMSLEKRFQGSAA 727  
 Db 443 ----- 442  
 QY 728 FVASTLMEYGVGPSTPESLLKEAIIHVISCGYEDKSEMGTEIGWTYGSVTEIDILGFKM 787  
 Db 443 -----LDEIGI-----AVE-----TVTEAHTSLRL 464  
 QY 788 HARGWRSYVCMKPKPAFKSAPINLSDRNLQVLRWALGSVEILFSRHCPGLWYGGRLKF 847  
 Db 465 HRRGTTAYM--RIPQAGLATESLSAHIGQIRWARGAVOI-FRLDNPL---FGGLKL 518  
 QY 848 LERFAYINTTIYPLTSLPLVLCILPAICLLTGKFTIMPEISNLASLWIFIALFISIFA-TG 906  
 Db 519 AQLCYLNAHPHFLSGIPRLIFLTAPLAFLLHAYI-----IYAPALMIALEVIPRH 569  
 QY 907 ILEMWSCVGIDEWNRNEQWNVIGGISAHLEFAVFGQLLKVLGIDTN---FTVTSKANDE 963  
 Db 570 MVHASLNSKIQGKYRH-SFW--SEIYETVLAWYIAPPTLVALINPHKGFENVTA- 622  
 QY 964 EGDFAEALYMEFKWTTLLIPPTTILINNVGVVAGTSYAINSGYQSGWGLFGKLFPAFWIV 1023  
 Db 623 -GGVVEEKYVDW-VISRPYIFLVLLNLGVAAGV-----WRYYGVP----- 661  
 QY 1024 HLYFPLKGLMGRQNRPTTIVVWVLLASIFSLWV 1059  
 Db 662 -----ENETLTIV-----SLWV 675  
 RESULT 3  
 ACS2\_ACEXY STANDARD; PRT; 1596 AA.  
 ID ACS2\_ACEXY  
 AC Q59167;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cellulose synthase 2 [Includes: Cellulose synthase catalytic subunit  
 (UDP-forming) (EC 2.4.1.12); Cyclic di-GMP binding domain (Cellulose  
 synthase 2 regulatory domain)].  
 GN ACSAII.  
 OS Acetobacter xylinus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;  
 OC Gluconacetobacter.  
 OX NCBI\_TaxID=28448;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 23769;  
 RX MEDLINE=95394846; PubMed=7665515;  
 RA Saxena I.M., Brown R.M. Jr.;  
 FT "Identification of a second cellulose synthase gene (acsAII) in  
 Acetobacter xylinum.";  
 RL J. Bacteriol. 177:5276-5283(1995).  
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + ((1,4-beta-D-glucosyl))(N) - UDP  
 + ((1,4-beta-D-glucosyl))(N+1).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (potential).  
 CC -!- DOMAIN: There are two conserved domains in the globular part of  
 CC the catalytic subunit: the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC OXRRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -!- MISCELLANEOUS: It is not essential for cellulose production in

















Qy	696	IEGVGAGCDDEKSVLMSQMSLERFQCQRAAFVASTLMETVGGVQSQSPESLLKEAIHV	755
Db	293	-FYGVTDGND-----FWDATFCGSCAI-----LRREAIES	323
Qy	756	ISCGYEDKSEWGTEIGWIYGSVTEDLTGFKHMGWRSVYCMKPRPAFKGSAPINLSDR	815
Db	324	IG-----GFAVETVTEDAHTALRMQRGWTAYL--RIPVASGLATERLUTH	368
Qy	816	LNQVLRWALGSVEILLFSRCHPLWYGGGLKFLERFAINTIYPLSLPLVYCIIPAI	875
Db	369	IGQRMWAGMIOI--FRVDNPM--LGRGLQGRLCYLSAMTGFPAIRVIFLASPLA	424
Qy	876	CLLTGTFIMPETNSLAIWFIALFSLISFATGILEMRWS---GVGIDEWNRNEQFV-VIGG	931
Db	425	FLFAGQNI-----AAAPLAAVAYALPHMHFISIATAAKVNKGWR-YGFSESVYE	472
Qy	932	ISAHLEFAVFGQLKVLGADIDTFEYTSKA---NDEEGDFAELY-MFKWTTLLIPPTPI--	985
Db	473	TTMALFLRVRTIVTLFPPSGKFNFTTEKGVLEEEFOLDGATYNNIIFATIMMGLLIGL	532
Qy	986	--LII--NMVGVVAGTSYAINSGYOSWGPFGLKFJFAFWIVHLYIFLKLGL-MGRQNR	1038
Db	533	FELIVAFNOLDIARAYALLNCA-----WALISIIILFAAIAVGREYK	575

Dd	732 TTWALFLVRVTVITLLFPSPKGFNTEKGVLVEEFDLGATYNNIIIFATIMMGGLLIGL
Dd	743 TTTWALFLVRVTIVTLFPPSKGFNFTEKGVLVEEFDLGATYNNTIIPATIMMGGLLIGL
Qy	986 --LI--NMGVVVAGTSTYAINSGYSWGPGLFKGLFFAFAFWVIHLYPFLKGI-MGRQNR
Dd	533 FELIVRFNQLDVIARNAYLLNCA-----WALISLTILLFAAIAVGRETK

```

RESULT 11
BCSA_XANAC                                STANDARD;          PRT;    729 AA.
ID   BCSA_XANAC                            CDS             1..729
AC   P58932;
DT   DT 15-JUN-2002 (Rel. 41, Created)
DD   DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE   DE 15-JUN-2002 (Rel. 41, Last annotation update)
DI   Di Cellulose synthase catalytic subunit [UDP-forming] (EC 2.4.1.12).
GN   GN BCSA_OR XAC3518.
OS   OS Xanthomonas axonopodis (pv. citri).
OC   OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
CX   CX Xanthomonas.
NCBI NCBI_TaxId=92829;
RN   RN [1]
RP   RP SEQUENCE FROM N.A.
```

RC STRAIN=306 / ATCC 13902 / XV 101;  
RX MEDLINE=22022145; PubMed=12024217;  
RA Da Silva A.C.R., Ferro J.A., Reinach C.B., Farah C.S., Furlan L.R.,  
RA Quaglio R.B., Monteiro-Vitorello C.F., Van Sluys M.A., Almeida N.F.,  
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,  
RA Camarotte G., Cannavan F., Cardoso J., Chambergó F., Clapina L.P.,  
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,  
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,  
RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,  
RA Katsuyama A.K., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,  
RA Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,  
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,  
RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,  
RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,  
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,  
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,  
RA Setubal J.C., Kitajima J.P.;  
RT "Comparison of the genomes of two Xanthomonas pathogens with differing  
RT host specificities.";  
EL Nature 417:459-463(2002).  
CC -I- FUNCTION: Catalytic subunit of cellulose synthase. It polymerizes  
CC uridine 5'-diphosphate glucose to cellulose, which is produced as  
CC an extracellular component for mechanical and chemical protection  
CC (by similarity).  
CC -I- CATALYTIC ACTIVITY: UDP-glucose + {(1,4-beta-D-glucosyl)}(N) = UDP  
CC + {(1,4-beta-D-glucosyl)}(N+1).  
CC -I- COFACTOR: Magnesium (By similarity).  
CC -I- ENZYME REGULATION: Activated by bis-(3'-5'), cyclic diguanylic acid  
CC (c-di-GMP) (By similarity).  
CC -I- PATHWAY: Bacterial cellulose biosynthesis.  
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (potential).  
CC -I- DOMAIN: There are two conserved domains in the globular part of

CC the protein: the N-terminal domain (domain A) contains  
CC the conserved DXD motif and is possibly involved in catalysis and  
CC substrate binding. The C-terminal domain (domain B) contains the  
CC QXXRW motif and is present only in processive glycosyl  
CC transferases. It could be involved in the processivity function of  
CC the enzyme, possibly required for holding the growing glycan chain  
CC in the active site.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL; AE012000; AAM38361.1; -  
CC  
CC Cellulose biosynthesis; Transferase; Glycosyltransferase;  
CC Transmembrane; Inner membrane.  
CC FT DOMAIN 151 244 CATALYTIC SUBDOMAIN A.  
CC FT DOMAIN 321 381 CATALYTIC SUBDOMAIN B.  
CC FT TRANSMEM 30 50 POTENTIAL.  
CC FT TRANSMEM 110 130 POTENTIAL.  
CC FT TRANSMEM 171 191 POTENTIAL.  
CC FT TRANSMEM 405 425 POTENTIAL.  
CC FT TRANSMEM 427 447 POTENTIAL.  
CC FT TRANSMEM 520 540 POTENTIAL.  
CC FT TRANSMEM 549 569 POTENTIAL.  
CC FT TRANSMEM 610 630 POTENTIAL.  
CC FT ACT\_SITE 193 193 POTENTIAL.  
CC FT ACT\_SITE 337 337 POTENTIAL.  
CC FT SITE 240 240 SUBSTRATE BINDING (POTENTIAL).  
CC FT SITE 242 242 SUBSTRATE BINDING (POTENTIAL).  
CC SQ SEQUENCE 729 AA; B9C08BB995E795B1 CRC64;

Query Match 4.0%; Score 230; DB 1; Length 729;  
Best Local Similarity 19.2%; Pred. No. 2.3e-011;  
Matches 149; Conservative 94; Mismatches 211; Indels 324; Gaps 31;  
  
QY 278 VIVRLVLSI-----FLHYRLTNP--VRNA--YPLWLLSVICELWPAFSLWILQFPKWF 328  
DB 77 VVLMGMSLAVSRYIWNRTQTMGVGSADVDFILGLGLGAELYAFVILVGFQVLM 136  
QY 329 PINRTYLDRLALYDRSGESQAAVDIFVSTVDPLKEPPIVANTVLSILAVDPVK 388  
DB 137 PLNRKP-----VPLPADQLRPS---VDVFIPTYN---EPLSVVVRTVLAASVIDWPAGK 185  
QY 389 VSCYVSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPEF-YFCOKIDYLLKDKVQP 447  
DB 186 YTHLLDGG-----RDEFRAPCAEV----- 206  
QY 448 SPVKDRAMKREYEEFKIRINALVSKALKVPEEGWIMODGTPWPGNNTDRHPGMIOVFLG 507  
DB 207 ----- 206  
QY 508 HSGGLDTSGNELPLRVYVSRKRCFQHHKAGAMNALVRYSAVLNCOYMLNLCDDHYI 567  
DB 207 ---GIN-----YVTRNNA---HAKAGINAAALK-----KCSGDYVAIFDCDH-I 244  
QY 568 NNSKAVREAM-CFLMDPNLGPQVCYVQPPQRF---DGIDRN-----DRYANENTVFFDINL 619  
DB 245 PTRSPQLQVAMGFWLHDTKL-----ALVQMPHYFFSPDPPEPRLNLDTHGKVPNEGELFYGLLQ 300  
QY 620 RGLDGIQGVVYGVGVNFRATVGYEPPITAKKPGFLASLCGGKKKASKSKRSSDKKK 679  
DB 301 DGNQDNWATFCGSCAVIKRTA----- 322  
QY 680 SNKHVDSSVPFNLEIDIEEGVAGAGFDEKSVLMSQMSLEKRFQSGAFAFVASTLMYEGGV 739  
DB 323 -----LEEYGV 329  
QY 740 PQSPTPELLKEATHIVISCGYEDKSEWGTETGWTGYSVTEDILYGFPMHARGWRSVCM 799

DB 330 -----AVE-----  
QY 800 KRPAKFSAPINLSDRNQVLR  
DB 357 --PQAAGLATESLSGHVQRI  
QY 860 PLTSLPLVYCIUPLAICLLTGR  
DB 411 FFYGVPRHIIYLTAPLFFGA  
QY 919 EWMNEQFWIGGISAHFAV  
DB 471 EYVETTLAWI-----E  
QY 973 FKWTLLPIPTTILIIINMGV  
DB 521 F-----LLLLNVGM  
  
RESULT 12  
ACSL\_ACEXY STANDARD;  
ID ACSL\_ACEXY STANDARD;  
DT 01-MAY-1991 (Rel. 18, Creation update)  
DT 15-JUN-2002 (Rel. 41, Last  
DE Cellulose synthase 1 (Inclu  
DE [UDP-forming] (RC 2.4.1.12)  
DE synthase 1 regulatory domain  
GN ACSAB OR ACSA OR ACSB.  
OS Acetobacter xylinus.  
OC Bacteria; Proteobacteria;  
OC Gluconacetobacter.  
OX NCBI\_TaxID=28448;  
RN [1]  
RP SEQUENCE FROM N.A., AND SE  
RC STRAIN=ATCC 53582;  
RX MEDLINE=91346705; PubMed=2  
RA Saxena I.M., Lin F.C., Bro  
RT "Cloning and sequencing of  
RT gene of Acetobacter xylinu  
RL Plant Mol. Biol. 15:673-68  
RN [2]  
RP SEQUENCE FROM N.A., AND SE  
RC STRAIN=ATCC 53582;  
RX MEDLINE=91322509; PubMed=1  
RA Saxena I.M., Lin F.C., Bro  
RT "Identification of a new ge  
RT in Acetobacter xylinum".  
RL Plant Mol. Biol. 16:947-95  
RN [3]  
RP REVISIONS.  
RC STRAIN=ATCC 53582;  
RX PubMed=8083166;  
RA Saxena I.M., Kudlicka K.,  
RT "Characterization of genes  
RT operon) of Acetobacter xyl  
RT crystallization.";  
RL J. Bacteriol. 176:5735-575  
RN [4]  
RP SEQUENCE OF 1-8 FROM N.A.  
RC STRAIN=ATCC 23769;  
RX MEDLINE=94131945; PubMed=8  
RA Standal R., Iversen T.-G.,  
RA Valla S.;  
RT "A new gene required for c  
RT cellulolytic activity in A  
RT bcs operon.";  
RL J. Bacteriol. 176:665-672(  
RN [5]  
RP FUNCTION.  
RC STRAIN=ATCC 53582;

RX PubMed-2138620;  
 RA Lin F.C., Brown R.M. Jr., Drake R.R. Jr., Haley B.E.:  
 RT "Identification of the uridine 5'-diphosphoglucose (UDP-Glc) binding  
 RT subunit of cellulose synthase in *Acetobacter xylinum* using the  
 RT photoaffinity probe 5-azido-UDP-Glc.";  
 RL J. Biol. Chem. 265:4782-4784(1990).  
 [6]  
 RP 3D-STRUCTURE MODELING, AND MUTAGENESIS OF D-188; D-189; D-236; D-333;  
 RP Q-369; R-370 AND R-372.  
 RC STRAIN-ATCC 23769;  
 RA PubMed-11430986;  
 RX Saxena I.M., Brown R.M. Jr., Dandekar T.:  
 RA "Structure-function characterization of cellulose synthase:  
 RT relationship to other glycosyltransferases.";  
 RL Phytochemistry 57:1135-1148(2001).  
 [7]  
 RN REVIEW ON DOMAIN ARCHITECTURE.  
 RP PubMed-7883697;  
 RX Saxena I.M., Brown R.M. Jr., Fevre M., Geremia R.A., Henrissat B.:  
 RA Multidomain architecture of beta-glycosyl transferases: implications  
 RT for mechanism of action.";  
 RL J. Bacteriol. 177:1419-1424(1995).  
 CC -!- FUNCTION: Bifunctional protein comprised of a catalytic subunit  
 CC and a regulatory subunit. The catalytic subunit of cellulose  
 CC synthase polymerizes uridine 5'-diphosphate glucose to cellulose  
 CC in a processive way. The thick cellulosic mats generated by this  
 CC enzyme probably provide a specialized protective environment to  
 CC the bacterium. The regulatory subunit binds bis-(3'-5') cyclic  
 CC diguanylic acid (c-di-GMP).  
 CC -!- CATALYTIC ACTIVITY: UDP-glucose + [(1.4)-beta-D-glucosyl](N) = UDP  
 CC + [(1.4)-beta-D-glucosyl](N+1).  
 CC -!- COFACTOR: Magnesium (By similarity).  
 CC -!- ENZYME REGULATION: Activated by c-di-GMP.  
 CC -!- PATHWAY: Bacterial cellulose biosynthesis.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -!- DOMAIN: There are two conserved domains in the globular part of  
 CC the catalytic subunit; the N-terminal domain (domain A) contains  
 CC the conserved DXD motif and is possibly involved in catalysis and  
 CC substrate binding. The C-terminal domain (domain B) contains the  
 CC QXXRW motif and is present only in processive glycosyl  
 CC transferases. It could be involved in the processivity function of  
 CC the enzyme, possibly required for holding the growing glycan chain  
 CC in the active site.  
 CC -!- INDUCTION: Cellulose is produced at a linear rate with respect to  
 CC cell growth when O12 is present.  
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE  
 CC GLYCOSYLTRANSFERASE FAMILY 2.  
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE ACSB/BCSB  
 CC FAMILY.  
 CC -!- CAUTION: Was originally (Ref.1) thought to be two separate ORFs  
 CC named acsa and acsb, due to a frameshift in position 678.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 -----  
 CC EMBL; X54676; CAA38487.1; ALT\_FRAME.  
 CC DR EMBL; X54676; CAA38488.1; ALT\_FRAME.  
 CC DR EMBL; M96060; AAA16971.1; .  
 CC DR PIR; S14437; S14437.  
 CC DR PIR; S13732; S13732.  
 CC DR InterPro; IPR003919; Cellsynth\_A.  
 CC DR InterPro; IPR001173; Glycos\_transf\_2.  
 CC DR Pfam; PF00535; Glycos\_transf\_2; 1.  
 CC DR PRINTS; PRO1439; CELLSYNTHASEA.  
 CC Cellulose biosynthesis; Transferase; Glycosyltransferase;  
 KW Transmembrane; Inner membrane.  
 KW DOMAIN 1 741 CATALYTIC.

FT	DOMAIN	742	1550		
FT	DOMAIN	147	240		CYCLIC DI-GMP BINDING.
FT	DOMAIN	317	377		CATALYTIC SUBDOMAIN A.
FT	TRANSMEM	26	46		POTENTIAL.
FT	TRANSMEM	47	67		POTENTIAL.
FT	TRANSMEM	106	126		POTENTIAL.
FT	TRANSMEM	398	418		POTENTIAL.
FT	TRANSMEM	423	443		POTENTIAL.
FT	TRANSMEM	468	488		POTENTIAL.
FT	TRANSMEM	507	527		POTENTIAL.
FT	TRANSMEM	547	567		POTENTIAL.
FT	TRANSMEM	1513	1533		POTENTIAL.
FT	ACT_SITE	333	333		POTENTIAL.
FT	SITE	236	236		SUBSTRATE BINDING (POTENTIAL).
FT	MUTAGEN	188	188		D->N: DECREASE IN ACTIVITY.
FT	MUTAGEN	188	188		D->P: LOSS OF ACTIVITY.
FT	MUTAGEN	189	189		D->Y: LOSS OF ACTIVITY.
FT	MUTAGEN	236	236		D->I: LOSS OF ACTIVITY.
FT	MUTAGEN	333	333		D->R: LOSS OF ACTIVITY.
FT	MUTAGEN	369	369		Q->M: LOSS OF ACTIVITY.
FT	MUTAGEN	370	370		R->P: LOSS OF ACTIVITY.
FT	MUTAGEN	370	370		R->Q: DECREASE IN ACTIVITY.
FT	MUTAGEN	372	372		R->A: LOSS OF ACTIVITY.
SQ	SEQUENCE	1550 AA;	168161 MW;	63AB8952BC39E961	CRC64;

Query Match 3.6%; Score 207; DB 1; Length 1550;  
 Best Local Similarity 19.6%; Pred. No. 3e-06;  
 Matches 154; Conservative 88; Mismatches 245; Indels 298; Gaps 29;

QY	280	VLRVVLVLSIFHLRTPVNRNAYPLWL	-----LSVTCETWFLSWITDQFPRKWFPI	NRE 333
DB	80	VLSGLVSLRYLTWRLTETL	--SPDTWLGGLTMLVAELVYALMMFLSYFQIAPLHRA	137
QY	334	TYLDRLALYDRREGEPSQLAAVDIFVSTVDPLKEPPIVTANTVLSILAVDPYDVKVCYV	393	
DB	138	P-----LPL-----PPNPDEWPTVDIFVPTYN	--EELSIVRLTVLGLSLGIDMPPEKVRVHI	186
QY	394	SDGASMLTFDALABTSEFARKKVPFVKYDIEPRAPFVFCOKIDYLDKQVPSFVKDR	453	
DB	187	LDGG-----	-----RREF-----	195
QY	454	RAMKREYEEFKIRINALVSKALKVPEGWIMQDTPPGNTRDHPGMIOVFLGHSGGLD	513	
DB	196	-----	-----AAFAECGA--	204
QY	514	TEGNELPRLVYVSRKRPFGQHHKAGAMNALVRVSAVLNGOYMLNCDCHVNNKAV	573	
DB	205	-----NYIART-----NEHAKAGNLNYAIG	-----HTGDYILIFDCDH-VPTRAFL	246
QY	574	REAMCFMDPNLGPQVCYVQFPQDFGIDRNDRYANRNTVFFDINLGLDGIQGPVYVGT	633	
DB	247	QLTGMWVE--DPKIALMQTPHHFYSPD	-----	272
QY	634	GCVENRTALYGEPPPIKAKKPGFLASLGGKKKASKSKRSSDKKSKNKHVDSSVPVFN	693	
DB	273	--PFORNLSAGVTRP-----	-----PEGNL	290
QY	694	EDIEEGVAGDFDEKSVLMSQMSLEKRFQSGAFAVATLMVEYGGVPSQSPESLKEAI	753	
DB	291	---FYGVVDGND-----FWDATEFCGSCAI	-----LRRTAI	319
QY	754	HVISCYEDKSEWGTGIGNYGVSTEDILTGFKMHASGWSVYCMKPRKPAKSGAPINLS	813	
DB	320	EQIG-----GFATQTVTEDAHTAKMRGLGWSTAYL	--RIPLAGGLATERLI	364
QY	814	DLRLNQRLWALGSVEILFSRHCPLWYGGRLKFLERFAYINTIYPLTSLPLCYILP	873	
DB	365	LHIGORVWARGLQI-FRIDNPL--FGRGLSNGORLCYLSAMTSEFAVPVPLFSSP	420	
QY	874	ACLTLTGKFIEMPEISNLSIASIFALFSIFATGILEMRWSCVG	-----IDENWRNEQW-V	928



QY 1034 GRONR 1038  
 Db 567 GRETK 571

RESULT 14  
 SUBF\_BACSU  
 ID SUBF\_BACSU STANDARD; PRT; 1433 AA.  
 AC P16397;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Bacillopeptidase F precursor (EC 3.4.21.-) (Esterase) (RP-I protease)  
 DE (90 kDa serine proteinase).  
 DE BPR OR BPF.  
 GN  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 195-222.  
 RX MEDLINE=90170864; PubMed=2106512;  
 RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,  
 RA Pero J.;  
 RT "Bacillopeptidase F of Bacillus subtilis: purification of the protein  
 RT and cloning of the gene.";  
 RL J. Bacteriol. 172:1470-1477(1990).  
 RN [2]  
 RP REVISIONS.  
 RC STRAIN=168;  
 RX MEDLINE=90368623; PubMed=2118514;  
 RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,  
 RA Pero J.;  
 RL J. Bacteriol. 172:5520-5521(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90216713; PubMed=2108961;  
 RA Wu X.-C., Nathoo S., Pang A.S.-H., Carne T., Wang S.-L.;  
 RT "Cloning, genetic organization, and characterization of a structural  
 RT gene encoding bacillopeptidase F from Bacillus subtilis.";  
 RL J. Biol. Chem. 265:6845-6850(1990).  
 RN [4]  
 RP SEQUENCE OF 1-211 FROM N.A.  
 RX MEDLINE=89008108; PubMed=3139638;  
 RA Beall B., Lowe M., Lutkenhaus J.;  
 RT "Cloning and characterization of Bacillus subtilis homologs of  
 RT Escherichia coli cell division genes ftsZ and ftsA.";  
 RL J. Bacteriol. 170:4855-4864(1988).  
 RN [5]  
 RP SEQUENCE OF 1410-1433 FROM N.A.  
 RC STRAIN=168 / Marburg;  
 RX MEDLINE=90174995; PubMed=2106671;  
 RA Masuda E.S., Anaguchi H., Sato T., Takeuchi M., Kobayashi Y.;  
 RT "Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus  
 RT subtilis.";  
 RL Nucleic Acids Res. 18:657-657(1990).  
 RN [6]  
 RP SEQUENCE OF 195-219.  
 RC STRAIN=Natto 16;  
 RX Kato T., Yamagata Y., Arai T., Ichishima E.;  
 RT "Purification of a new extracellular 90-kDa serine proteinase with  
 RT isoelectric point of 3.9 from Bacillus subtilis (natto) and  
 RT elucidation of its distinct mode of action.";  
 RL Biosci. Biotechnol. Biochem. 56:1166-1168(1992).  
 CC -!- SUBCELLULAR LOCATION: Secreted.  
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC  
 DR EMBL; M29035; AAA62679.1; -;  
 DR EMBL; J05400; AAA83362.1; -;  
 DR EMBL; M22630; AAA2458.1; -;  
 DR EMBL; X17344; CAA35224.1; -;  
 DR EMBL; Z99111; CAA13403.1; -;  
 DR EMBL; Z99112; CAA13404.1; -;  
 DR PIR; A35131; A35131.  
 DR PIR; A35750; A35750.  
 DR PIR; A36734; A36734.  
 DR HSP; P00782; 2SBT.  
 DR MEROPS; S08.017; -;  
 DR SubtilList; BG10233; bpr.  
 DR InterPro; IPR00209; Peptidase\_S8.  
 DR Pfam; PF00082; Peptidase\_S8; 1.  
 DR PRINTS; PR00723; SUBTILISIN.  
 DR PROSITE; PS00136; SUBTILASE\_ASP; FALSE\_NEG.  
 DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
 DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
 KW Hydrolase; Serine protease; Zymogen; Signal; Complete proteome.  
 FT SIGNAL 1 30  
 FT PROPEP 31 194 POTENTIAL.  
 FT CHAIN 195 755 BACILLOPEPTIDASE F.  
 FT PROPEP 756 1433 POTENTIAL.  
 FT ACT\_SITE 227 227 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 274 274 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT ACT\_SITE 452 452 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT CONFLICT 219 219 T -> A (IN REF. 6).  
 FT CONFLICT 393 393 A -> V (IN REF. 3).  
 FT CONFLICT 829 834 KHONKA -> N (IN REF. 3).  
 FT CONFLICT 836 841 OPOLV -> RTRYS (IN REF. 3).  
 FT CONFLICT 844 852 AQSVYETG -> FCRSRHKS (IN REF. 3).  
 FT CONFLICT 853 1433 MISSING (IN REF. 3).  
 SQ SEQUENCE 1433 AA; 154577 MW; 98DF6846897807C9 CRC64;

Query Match 2.1%; Score 119; DB 1; Length 1433;  
 Best Local Similarity 19.8%; Pred. No. 4.6;  
 Matches 184; Conservative 96; Mismatches 314; Indels 334; Gaps 49;

QY 13 HG---AGDVCQICADGLGTLTLDGVFTACDVCPRPVCPCYEHER----- 54  
 Db 503 HGLVNAFADAVSAVTDGLGKA-EGQVSVEGDQDEPPV---YQHEKVTVEYEGGSLPTLT 557  
 QY 55 KEGTQACLOCKTKYKRHRG-----SPAIRGEEGDDTDADDGSDFNYPASGTEQKOKIA 108  
 Db 558 AEDNVSVTSVKLSYKLDQGEWTEITAKRISGDHLKGT-----YQAEIPDIKGTKLS 608  
 QY 109 DMRSWRMNTGGSGNVGHPKYD-----SGEIGLSKYDSGEIPRGYVSVTNSQMSGEIPG 163  
 Db 609 YK---WMTHDFGGHVSSDVIDVTVPKPSITAGY-KQDFETAPGGWASGTNNNWEVGPS 664  
 QY 164 ASPDH-----HMSPGTGNISRRAPFYVNVHSPNPSREFSGSI----- 200  
 Db 665 TGPNTAASGEKYGTNLGTGNVANSANMLV-MPIKAPDSGSLFLQFKSHNLEDDFDY 722  
 QY 201 -----GNVAMK-----ERVD--GWKMKQDKGAIPMTNGTSTIAPSGRA 236  
 Db 723 GYFVFLPEGEKNWEQAGVINGKTSWTDEEIDLKAYKQNTQVMFNQSDSIA-KEGWY 781  
 QY 237 ATDI---DASTEYNMEDALLNDETQPLSRKVPVIASSKINPYRMVIVLRVLVLSIFLHYR 293  
 Db 782 IDDVVLSDKSAGTKVKKNKLGVEKPSGKQKKKPNPKAKPSANTAVKH-----QNK 833  
 QY 294 LTNPNRAYPLWLLSVICEIWFALSWILDQPKWFPINRETYLDRLALRY---DREGEPS 350  
 Db 834 ATQP--QVLPLKAQSVVVE-----TGKSTYSQSTGQVTLKHKGADYT 874  
 QY 351 QLAADVDFVSTVD--PLKEPPIVTANTVLSLT-----ADVDPVKVSCYSD 396  
 Db 875 LMAEYGTQSKTKVSLKTDQTTQANFTLEMKKGTGKGVINKTGETPVGASVYVED 934  
 QY 397 GA---SMLTFDA---LAETSEFARKWVPVKKYDIEPRAPEFYFCQKIDYLDKDKVQPSFVK 451





Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	5423.5	93.9	1079	10	Q9LTL1	Q9L111	zea mays (m
2	5421.1	93.8	1077	10	Q9LTL6	Q9L116	zea mays (m
3	5143	89.0	1076	10	Q9LTL5	Q9L115	zea mays (m
4	4577.5	79.2	1067	10	Q9XGX6	Q9X9X6	gossypium h
5	4530	78.4	1065	10	Q9FHK6	Q9FNK6	arabidopsis
6	4507	78.0	1065	10	Q48948	Q48948	arabidopsis
7	3993.5	69.1	1074	10	Q9LTL8	Q9L118	zea mays (m
8	3981	68.9	1075	10	Q9LTL9	Q9L119	zea mays (m
9	3963.5	68.6	1081	10	Q48946	Q48946	arabidopsis
10	3844.5	66.5	1094	10	Q9LTL2	Q9L112	zea mays (m
11	3820	66.1	1059	10	Q9LTL4	Q9L114	zea mays (m
12	3818	66.1	1086	10	Q9LTL3	Q9L113	zea mays (m
13	3801	65.8	1026	10	Q9SWM6	Q9SWM6	arabidopsis
14	3795	65.7	1026	10	Q9XHP6	Q9XHP6	arabidopsis
15	3778.5	65.4	1065	10	Q9SKJ5	Q9SKJ5	arabidopsis
16	3746	64.8	1043	10	Q9FNC3	Q9FNC3	arabidopsis

Db	61	CPOCKNKKRHKSGPAIRGEEDDTDADSDAFNYPASGNDQKQKIADMRSRWNRNAGG	120
Qy	121	SGNVGHPKYDSGEIGLSKYSDEIRGYPVSTNSOMSGEIPGASPDHMMSPGTGNISR	180
Db	121	SGDVGPKYDSGEIGLTKYDSGEIPGYPSTNSOISGEIPGASPDHMMSPGTGNIGR	180
Qy	181	APFPYVNHSPNSREFSGSIGNVANKERVDGWMKQDKGAIPTWNTGTSIAPSEGRAATI	240
Db	181	APFPYVNHSSNSREFSGSGVNVANKERVDGWMKQDKCTIPMTNGTSIAPSEGRGVGI	240
Qy	241	DASTEYNMEDALLNDETROPISRKVPDIASKINPYRMVIVLRLVLSIFLHYRLTNPVNR	300
Db	241	DASTDYNMEDALLNDETROPISRKVPDIASKINPYRMVIVLRLVLSIFLHYRLTNPVNR	300
Qy	301	AYPLWLLSVICETWFAWLLDQFPKWFPIINRETYLDRLALRYDREGESQLAADVFS	360
Db	301	AYPLWLLSVICETWFAWLLDQFPKWFPIINRETYLDRLALRYDREGESQLAADVFS	360
Qy	361	TVDPLKEPPIVANTVISTILAVDYPVDKYSVSDDGASMLTFDALAETSEFARKWVPFV	420
Db	361	TVDPMKEPPLVANTVLSILAVDYPVDKYSVSDDGAAAMLTFDALAETSEFARKWVPFV	420
Qy	421	KYDIEPRAPEFYFCOKIDYLDKQVOPSVFKDRAMKREYEFEKIRINALVSKALVPEE	480
Db	421	KYINIEPRAPEWYFSQKIDYLDKQVOPSVFKDRAMKREYEFEKIRVNGLVAKAQVPEE	480
Qy	481	GMTQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSRKRPFGFHHKAG	540
Db	481	GMTQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSRKRPFGFHHKAG	540
Qy	541	AMNALVRVSAVLINGOYMLNLDCHYINNSKAVREAMCFMLDPNLGPQVCYVQFPQFDDG	600
Db	541	AMNALVRVSAVLINGOYMLNLDCHYINNSKALREAMCFMLDPNLGRSVCYVQFPQFDDG	600
Qy	601	IDRNDRYANRNTVFFDINLRGLDIOGYPVYVGTGCFNRTAIYGYEPPKAKKPGFLASL	660
Db	601	IDRNDRYANRNTVFFDINLRGLDIOGYPVYVGTGCFNRTALYGPPIKQKKGFLSSL	660
Qy	661	CGGKKKASKSKRSKSKSNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLEK	720
Db	661	CGGRKKGSKSKK-GSDKKSQKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLEK	719
Qy	721	RFQSAAFVASTLMEYGGVQPSSTPSLLKEAITHVISCYEDKSEWGTGIGWYGSVTE	780
Db	720	RFQSAAFVASTLMEYGGVQPSATPSLLKEAITHVISCYEDKIEWGTGIGWYGSVTE	779
Qy	781	ILTGFKMHARGWSVCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPLWYG	840
Db	780	ILTGFKMHARGWSVCMKRPAPKGSAPINLSDRNLQVLRWALGVSVEILFSRHCPLWYG	839
Qy	841	YGRKLFLERFAYINTIYPLTSLPLLYCILLPAICLLTGKFTMPETISNLASLWFTALFL	900
Db	840	YGRKLFLERFAYINTIYPLTSLPLLYCILLPAICLLTGKFTIPEISNPFASLWFTSLFI	899
Qy	901	SIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA	960
Db	900	SIFATGILEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFGQLLKVLAGIDTNTFTVTSKA	959
Qy	961	NDEGDFAEILYMKFTWTLPIPTTILLINMGVGVAGTSYAINSGYQSWGFLFKLFFAFW	1020
Db	960	SDEGDFAEILYMKFTWTLPIPTTILLINMGVGVAGTSYAINSGYQSWGFLFKLFFAFW	1019
Qy	1021	VIVHLVPLKGLMGONRPTIIVVWAILLASIFSLWLWVDPFTTRLAGPNQTCGINC	1080
Db	1020	VIVHLVPLKGLMGONRPTIIVVWAILLASIFSLWLWVDPFTTRLAGPNQTCGINC	1079
RESULT 2			
Q9LLI6		PRELIMINARY; PRT; 1077 AA.	
AC	Q9LLI6;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		

DT	01-JUN-2002 (TrEMBLrel. 21, annotation update)		
DE	Cellulose synthase-4.		
GN	CESA-4.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;		
OC	Panicoidae; Andropogoneae;		
OX	NCBI_TaxID=4577;		
[1]	SEQUENCE FROM N.A.		
RN	MEDLINE=20398328; PubMed=10770;		
RX	Holland N., Holland D., Helander P.;		
RA	Xoconostle-Cazares B., Delmar P.;		
RT	"A comparative analysis of plant cellulose synthase (Cesa) gene family.";		
RL	Plant Physiol. 123:1313-1322 (2000).		
DR	EMBL; AF200528; AAF89964.1; InterPro: IPR005150; Cellulose synthase; Pfam: PF03552; Cellulose_sy SMART; SM00184; RING; 1.		
DR	SMART; SM00184; RING; 1.		
SQ	SEQUENCE 1077 AA; 120501		
Query Match	93.8%;		
Best Local Similarity	92.4%;		
Matches 998; Conservative			
Qy	1	MDGDADALSKRGAGDVCQIC	60
Db	1	MEDGADGVKSGRGGGQVCQIC	60
Qy	61	CLOCKTKYKRHRGSPAIRGEEG	120
Db	61	CPOCKTKYKRHKGSPAIRGEEG	118
Qy	121	SGNVGHPKYDSGEIGLSKYDSG	180
Db	119	SGDVGPKYDSGEIGLTKYDSG	178
Qy	181	APFPYVNHSPNSREFSGSIGN	240
Db	179	APFPYVNHSPNSREFSGSIGN	238
Qy	241	DASTEYNMEDALLNDETROPISL	300
Db	239	DASTDYNMEDALLNDETROPISL	298
Qy	301	AYPLWLLSVICETWFAWLLDQFPKWFPIINRETYLDRLALRYDREGESQLAADVFS	360
Db	299	AYPLWLLSVICETWFAWLLDQFPKWFPIINRETYLDRLALRYDREGESQLAADVFS	358
Qy	361	TVDPLKEPPIVANTVISTILAVDYPVDKYSVSDDGASMLTFDALAETSEFARKWVPFV	420
Db	359	TVDPMKEPPLVANTVLSILAVDYPVDKYSVSDDGAAAMLTFDALAETSEFARKWVPFV	418
Qy	421	KYDIEPRAPEFYFCOKIDYLR	480
Db	419	KYINIEPRAPEWYFSQKIDYLR	478
Qy	481	GMTQDGTWPNGNTRDHPGMI	540
Db	479	GMTQDGTWPNGNTRDHPGMI	538
Qy	541	AMNALVRVSAVLINGOYMLNLD	600
Db	539	AMNALVRVSAVLINGOYMLNLD	598
Qy	601	IDRNDRYANRNTVFFDINLRGI	660
Db	599	IDRNDRYANRNTVFFDINLRGI	658
Qy	661	CGGKKKASKSKRSKSKSNKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLEK	720

```
Db 659 CGGRKKASKSK -GSDKKKSQKHVDSSVPVFNLEDIEBVGECAGDFDDEKSLMSQMSLEK 717
QY 721 RFQSAFVASTLMEYGGVQSSPTESLLKKAHIVISGVEDKSEWTEIGWIYGSVTD 780
Db 718 RFQSAFVASTLMEYGGVQSSPTESLLKKAHIVISGVEDKSEWTEIGWIYGSVTD 777
QY 781 ILTGFKMHARGWRSVYCMKPRAPKGSAPINLSDRLNOVLWALSGVEILFSRHCPLWYG 840
Db 778 ILTGFKMHARGWRSVYCMKPRAPKGSAPINLSDRLNOVLWALSGVEILFSRHCPLWYG 837
QY 841 YGRLKFLERFAYINTTIYPLTSLPLLYCYLPAICLTGKFIKPEISNLASIWFIALFL 900
Db 838 YGRLKFLERFAYINTTIYPLTSLPLLYCYLPAICLTGKFIKPEISNLASIWFIALFL 897
QY 901 SIFATGILEMRWSGVGIDEMWNEQFWVIGGISAHFLFAVFGQLLKLVLGIDTNTFTVSKA 960
Db 898 SIFATGILEMRWSGVGIDEMWNEQFWVIGGISAHFLFAVFGQLLKLVLGIDTNTFTVSKA 957
QY 961 NDEEGDFAELYMFKWTLLIPTTILINMVGWVAGTSYAINSGYQSGWGLFGLFFAFW 1020
Db 958 SDEGDFAELYMFKWTLLIPTTILINMVGWVAGTSYAINSGYQSGWGLFGLFFAFW 1017
QY 1021 VIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRLAGENIQTGINC 1080
Db 1018 VIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRVGTGDTQTGINC 1077

RESULT 3
Q9LLI5 PRELIMINARY; PRT; 1076 AA.
AC Q9LLI5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase-5.
GN CESA-5.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Panicoideae; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20398320; PubMed=10938350;
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,
RA Xoonostie-Cazares B., Delmer D.P.;
RT "A comparative analysis of the plant cellulose synthase (CesA) gene
family."
RL Plant Physiol. 123:1313-1324(2000).
DR EMBL; AF200529; AAF89965.1;
DR InterPro; IPR005150; Cellulose_synt.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1076 AA; 120724 MW; 41E7C20EDA54F27 CRC64;

Query Match 89.0%; Score 5143; DB 10; Length 1076;
Best Local Similarity 87.6%; Pred. NO. 0;
Matches 947; Conservative 69; Mismatches 59; Indels 6; Gaps 6;

QY 1 MDGDADALSKGRHAGDVCOICADGLGTLDGVFTACDVCRPPVPCPYEHERKEGTQA 60
Db 1 MDG-GDATNSGRHAGDVCOICGDGVGTAADGDLFTACDVCGFPVPCPYEHERKDGTA 59
QY 61 CLQCKTKYKRRHRSIPAIREGDDTDADDGSDFNYPASGTEDQKQADMRWSRMTGG 120
Db 60 CPCKTKYKRRHRSPPVHGEENEDVDADDVSDNYQASGNODQKQKTAERMLTWTNRSRG 119
QY 121 SGNVGHPKYDSGEIGLSKYDSGEIPRGVPSVTNSQMSGEIPGASPDHMMSPGTNISR 180
Db 120 S-DIGLAKYDSGEIGHGYDSGEIPRGVPSVTNSQMSGEIPGASPD-HMMSPGVNTGRR 177
QY 181 A-PFPVYVNHSPNPSREFSSIGNVAWKERVGDGWMKQDKGAIPWTNGTSTIAPSEGRATD 239
```

```
Db 178 GHQFPVNHSPNPSREFSSIGNVAWKERVGDGWMK -DKGAIPWTNGTSTIAPSEGRGVA 236
QY 240 IDASTEYNMEDALLNDETROPLSRKVPYPTASSKINPYRMVIVLRLVLSIFLHRLTNPVR 299
Db 237 IDASTDYNMEDALLNDETROPLSRKVPYPTASSKINPYRMVIVLRLVLSIFLHRLTNPVR 296
QY 300 NAYPLWLLSVICEIWFALSWILDQFPKWFPIINRETYIDRLALRYDREGEPSQLAAVDIFV 359
Db 297 NAYPLWLLSVICEIWFALSWILDQFPKWFPIINRETYIDRLALRYDREGEPSQLAPVDIFV 356
QY 360 STVDPKLEPPYIVTANTVLSILAVDPYDKVSCYVSDGASMLTFDALAEISEFARKWVPF 419
Db 357 STVDPKLEPPYIVTANTVLSILAVDPYDKVSCYVSDGASMLTFDALAEISEFARKWVPF 416
QY 420 VKKYDIEPRAPFEYFCQIDYLDKQVQSPFVKDRAMKREYEERKIRINALVSKALKVPE 479
Db 417 CKKYNIEPAPDPMYFAQIDYLDKQVQSPFVKDRAMKREYEERKIRINALVSKALKVPE 476
QY 480 EGWIMQDGTTPWGNTRDPHGMIVFLGHSGGLDTEGNEPRLVYVSRKRPQFQHKK 539
Db 477 EGWIMQDGTTPWGNTRDPHGMIVFLGHSGGLDTEGNEPRLVYVSRKRPQFQHKK 536
QY 540 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKAVREAMCFMDPNLGPQVYVQFQPF 599
Db 537 GAMNALVRYSAVLTNGQYMLNLDCHYINNSKALREAMCFMDPNLGNVYVQFQPF 596
QY 600 GIDRNDRYANRNTVFDDINLRLGDIQGPVYVGGCVFNRTAIYGYPPPKAKKPGFLAS 659
Db 597 GIDRNDRYANRNTVFDDINLRLGDIQGPVYVGGCVFNRTAIYGYPPPKAKKPGFLAS 656
QY 660 LCGGKKKSKSKSKSKSKKHVDSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLE 719
Db 657 LCGGKKKSKSKSKSKSHRHADSSVPVFNLEDIEEGVEGAGFDDKSVLMSQMSLE 715
QY 720 KRFQSAFVASTLMEYGGVQSSPTESLLKKAHIVISGVEDKSEWTEIGWIYGSVTE 779
Db 716 KRFQSSVEFVASTLMEYGGVQSSPTESLLKKAHIVISGVEDKSEWTEIGWIYGSVTE 775
QY 780 DILTGFKMHARGWRSVYCMKPRAPKGSAPINLSDRLNOVLWALSGVEILFSRHCPLWY 839
Db 776 DILTGFKMHARGWRSVYCMKPRAPKGSAPINLSDRLNOVLWALSGVEILFSRHCPLWY 835
QY 840 GYGRLKFLERFAYINTTIYPLTSLPLLYCYLPAICLTGKFIKPEISNLASIWFIALF 899
Db 836 GYGRLKFLERFAYINTTIYPLTSLPLLYCYLPAICLTGKFIKPEISNLASIWFIALF 895
QY 900 LSIFATGILEMRWSGVGIDEMWNEQFWVIGGISAHFLFAVFGQLLKLVLGIDTNTFTVSK 959
Db 896 LSIFATGILEMRWSGVGIDEMWNEQFWVIGGISAHFLFAVFGQLLKLVLGIDTNTFTVSK 955
QY 960 ANDEEGDFAELYMFKWTLLIPTTILINMVGWVAGTSYAINSGYQSGWGLFGLFFAF 1019
Db 956 ANDEEGDFAELYMFKWTLLIPTTILINMVGWVAGTSYAINSGYQSGWGLFGLFFAF 1015
QY 1020 WVIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRLAGENIQTG 1079
Db 1016 WVIVHLYPFLKGLMGQRNTPPTIIVYMAVLLASIFSLWVRVDPFTTRVGTGDIACGIN 1075
QY 1080 C 1080
Db 1076 C 1076

RESULT 4
Q9XGX6 PRELIMINARY; PRT; 1067 AA.
AC Q9XGX6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN CELA3.
```

OS Gossypium hirsutum (Upland cotton).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eucosids II; Malvales; Malvaceae; Gossypium.  
OX NCBI\_TaxID=3635;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. TEXAS MARKER-1; TISSUE=COTTON FIBER;  
RX MEDLINE=20098737; PubMed=10631273;  
RA Laosinchai W., Cui X., Brown R.M. Jr.;  
RT "A full length cDNA of cotton cellulose synthase has high homology  
with the Arabidopsis RSH1 gene and the cotton Cel1L gene (Accession  
No. AF200453) (PGR 00-002).";  
RL Plant Physiol. 122:291-291(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. TEXAS MARKER-1; TISSUE=COTTON FIBER;  
RA Kimura S., Laosinchai W., Itoh T., Cui X., Brown R.M. Jr.;  
RT "Immunogold Labeling of Rosette Terminal Cellulose Synthesizing  
Complexes in a Vascular Plant (Vigna angularis).";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DDBB databases.  
DR EMBL; AF150630; AAD39534.2; -;  
DR InterPro; IPR005150; Cellulose\_synt.  
DR Pfam; PF03552; Cellulose\_synt; 1.  
DR SMART; SM00184; RING; 1.  
SQ SEQUENCE 1067 AA; 119325 MW; 0BBA2ED00590F29C CRC64;  
Query Match 79.2%; Score 4577.5; DB 10; Length 1067;  
Best Local Similarity 78.6%; Pred. No. 0;  
Matches 853; Conservative 100; Mismatches 105; Indels 27; Gaps 11;  
QY 2 DGDADALKSGRHGAGDVQCICADGLCTLDGVTACDYPVCPVCEHERKEGTOAC 61  
DB 4 EGDIGG-KPMKNGGTCQICGDNVGNKTDGDFIACNCAFFVCPVCEYERKQNGQC 62  
QY 62 LOCKTKYKRRHSGPAIRGEGDDTDADDG-SDFNTPASGTEDQKQIADMRMSWRNMTGG 120  
DB 63 PCKTRKYKWKSGFAILGDRGTGDDAGDASDFY--SENQKQKQKLAERMGWNAKYR 120  
QY 121 SGNVGHPKYDSGEIGLSKYDSGRIPRGYVPSVTNSQ-MSGEIPGASPDHMHMSPTGNISR 179  
DB 121 GEDVGAPTYDX-----EISNHIPLTSGGEVSGEISASPERLSWASPGVAGG 169  
QY 180 RABPPYVNHSPNRF--SGSIGNVAKRVDGKWKQDKQGAIPMTNGTSIAPSGRAAT 238  
DB 170 KSSIRVV---DPVREFSGSGLVNVAKERVGKWKQEKNTVPMS--TCQATSE-RGLG 222  
QY 239 DIDASTEYNMEDALLNDETROPUSRRKVPITASSKINPYRMVIVLRLVLSIELHYRLTNPV 298  
DB 223 DIDASTDVLVDLSQNDQARQPLSRKVSSTKINPYRMVILRLVILCFLHYRLTNPV 282  
QY 299 RNAYPLWLSVCEIWFALSWILDQPKWFFPINRETYLDRALRYDREGEPSQLAAVDIF 358  
DB 283 PNAYALWLSVCEIWFALSWILDQPKWFFPINRETYLDRALRYDREGEPSQLAAVDIF 342  
QY 359 VSTVDPKPEPIVYANTVLSILAVDYPVKVSCYVSDDGASMLTFDALAETSEFAKQVP 418  
DB 343 VSTVDPKPEPIVYANTVLSILAVDYPVKVSCYVSDDGASMLTFDALAETSEFAKQVP 402  
QY 419 FVKYDIEPRAPFFYFCQIDYLDKQVPSFKVDRAMKREYEERKIRINALVSKALKVP 478  
DB 403 FCKYDIEPRAPFFYFAQIDYLDKQVPSFKVDRAMKREYEERKIRINALVSKALKVP 462  
QY 479 EEWIMODGTPWPGNTRDHPGMIOVFLGHSGLDPEGNELPRLVYVSREKRPQGHKK 538  
DB 463 EEWIMODGTPWPGNTRDHPGMIOVFLGHSGLDPEGNELPRLVYVSREKRPQGHKK 522  
QY 539 AGAMNALVRSVAVLTNGQYMLNDCDHYINNSKAVREAMCFMLDNLGQVQVYVFPQRF 598  
DB 523 AGAMNALVRSVAVLTNGAFLLNDCDHYINNSKALREAMCFMLDNLGQVQVYVFPQRF 582  
QY 599 DGDINDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGCTVFNRTAIYGYEPIKAK--KPGF 656

DB 583 DGDINDRNDYANRNTVFFDINLRGLDGIQGPVYVGTGCTVFNRTAIYGYEPIKAKHRKTQI 642  
QY 657 LASLCGG-KKASKSKKRSSDKKXVDSSVPVFNLEDIEEGVEGAGFDEKSLMSQ 715  
DB 643 LSSLCGGSKKSSKSSKRGSDKXVDSTVPVFNLEDIEEGVEGAGFDEKSLMSQ 702  
QY 716 MSLEKRRFGOSAAFVASTLMFYGSSTPESLLKEAIIHVISCYGYEDKSEWGTIGWIY 775  
DB 703 MSLEKRRFGOSAAFVASTLMFYGSSATPETLLKEAIIHVISCYGYEDKTDWGSSEIGWIY 762  
QY 776 SVTEDILTGPKMHARGWRSYVYPAFKGSAPINLSDRNLNOVLRWALGSVEILFSRHC 835  
DB 763 SVTEDILTGPKMHARGWRSYVYPAFKGSAPINLSDRNLNOVLRWALGSVEILFSRHC 822  
QY 836 PLWYGYGGRLKFLERFAYINTVSLPLLVYCIIPALCLLTGCKTFMPEISMLASIW 895  
DB 823 PIWYGYSGRLKWLERFAYINTVTAIPLLMYCTLPVAVCLLTNKFIPQISNLASIW 882  
QY 896 IALFLSIFATGILEMRWSGVGRNEQFWVIGGISAHLFAVFOGLLKVLAGIDINTFT 955  
DB 883 ISLFLSIFATGILKMKNGVGRNEQFWVIGGISAHLFAVFOGLLKVLAGIDINTFT 942  
QY 956 VTSKANDDEGDAEALYMEKWTPTTILINNVGVVAGTSYAINSGYSGWGLFGKL 1015  
DB 943 VTSKANDDEGDAEALYMEKWTPTTILINNVGVVAGTSYAINSGYSGWGLFGKL 1002  
QY 1016 FFAFWVIVHLVFFLKLGMGRQIVVNAVILLASTFSLLLWVRVDPFTTTLRAGNPQT 1075  
DB 1003 FFAFWVIVHLVFFLKLGMGRQIVVNAVILLASTFSLLLWVRVDPFTTTLRAGNPQT 1062  
QY 1076 CGINC 1080  
DB 1063 CGINC 1067  
RESULT 5  
Q9FHK6  
ID Q9FHK6 PRELIMINARY;  
AC Q9FHK6;  
DT 01-MAR-2001 (TrEMBLrel. 16) (date)  
DT 01-MAR-2001 (TrEMBLrel. 16) (sequence update)  
DE 01-JUN-2002 (TrEMBLrel. 21) (annotation update)  
DE Cellulose synthase catalytic subunit.  
OC Arabidopsis thaliana (Mouse eucosids II; Brassicales; Eukaryota; Viridiplantae; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RX MEDLINE=99397451; PubMed=10631273;  
RA Kaneko T., Katoh T., Sato M., Miyajima N., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX. Sequence features of the P1 and TAC clones."  
RL DNA Res. 6:183-195(1999).  
DR EMBL; AB018111; BAB09693.1;  
DR InterPro; IPR005150; Cellulose\_synt.  
DR Pfam; PF03552; Cellulose\_synt.  
DR SMART; SM00184; RING; 1.  
SQ SEQUENCE 1065 AA; 119665 MW; 3AA4714CE3C4D581 CRC64;  
Query Match 78.4%; Score 4530; DB 10; Length 1065;  
Best Local Similarity 78.6%; Pred. No. 0;  
Matches 839; Conservative 100; Mismatches 104; Indels 26; Gaps 8;  
QY 19 CQICADGLGTLDGVDVFTACDYPVCPVCEHERKEGTOACLOCKTKYKRRHSGPAIR 78  
DB 20 CQICSDNVGKTVDGDFVACDYPVCPVCEYERKQNGSCFQCKTRYKRLKSGPAIP 79

```
Qy 79 GEGDDTDADDGS--DFNYPASGTEDQKQIADRMRSWRMTGSGNVGHPKYDSGEIGLS 137
Db 137 GEGDDTDADDGS--DFNYPASGTEDQKQIADRMRSWRMTGSGNVGHPKYDSGEIGLS 137
Db 80 GKDDEGLADEGIVFNPY-----QKEKISERMLGWHLTRGKEEMGEPOYDK----- 127
Qy 138 KYDSEIPIRGVPSVTSNSQ--MSGEIPGASPDHMHMSTGNSIRAPFPY--VNHSPNPSRE 195
Db 195 KYDSEIPIRGVPSVTSNSQ--MSGEIPGASPDHMHMSTGNSIRAPFPY--VNHSPNPSRE 195
Db 128 -----EVSHNHLPLRTSRQDTSGEFSAASPERLSVSTTAGGRKRLYSDDVQSNRRIV 182
Qy 196 FSGSIVGNVANKERVGDGKMKQDKGAIPMTNGTSTIAPSEGRRAAIDIDASTEYNMEDALLND 255
Db 255 FSGSIVGNVANKERVGDGKMKQDKGAIPMTNGTSTIAPSEGRRAAIDIDASTEYNMEDALLND 255
Db 183 DPVGLGNVANKERVGDGKMKQDKGAIPMTNGTSTIAPSEGRRAAIDIDASTEYNMEDALLND 238
Qy 256 ETROPILSRKVPDIASSKINPYRMVIVLRLVLSIFLHYRLTNPNVRYNAYPLWLLSVICEIWF 315
Db 315 ETROPILSRKVPDIASSKINPYRMVIVLRLVLSIFLHYRLTNPNVRYNAYPLWLLSVICEIWF 315
Db 239 EAROPILSRKVPDIASSKINPYRMVIVLRLVLSIFLHYRLTNPNVRYNAYPLWLLSVICEIWF 298
Qy 316 ALSWILDOQFPKWPPIINRETYLDRLALRYDREGESQLAADVIFVSTVDPLKEPPIVTANT 375
Db 375 ALSWILDOQFPKWPPIINRETYLDRLALRYDREGESQLAADVIFVSTVDPLKEPPIVTANT 375
Db 299 ALSWILDOQFPKWPPIINRETYLDRLALRYDREGESQLAADVIFVSTVDPLKEPPIVTANT 358
Qy 376 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWPFVFKYDIEPRAPFAYFC 435
Db 435 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWPFVFKYDIEPRAPFAYFC 435
Db 359 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWPFVFKYDIEPRAPFAYFC 418
Qy 436 QKIDYLDKDKVQPSFVKDORRAMKREYEFEKIRINALVSKALKVPEEGWIMODGTPWPGNNT 495
Db 495 QKIDYLDKDKVQPSFVKDORRAMKREYEFEKIRINALVSKALKVPEEGWIMODGTPWPGNNT 495
Db 419 AKIDYLDKDKVQPSFVKDORRAMKREYEFEKIRINALVSKALKVPEEGWIMODGTPWPGNNT 478
Qy 496 RDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKRPQFQHHKAGAMNALVRYSVLTNG 555
Db 555 RDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKRPQFQHHKAGAMNALVRYSVLTNG 555
Db 479 GDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKRPQFQHHKAGAMNALVRYSVLTNG 538
Qy 556 QYMLNLCDDHYINNSKALREAMCFMDPNLGQVQYVQFPQREDGIDRNDYANRNTVFF 615
Db 615 QYMLNLCDDHYINNSKALREAMCFMDPNLGQVQYVQFPQREDGIDRNDYANRNTVFF 615
Db 539 PFILNLCDDHYINNSKALREAMCFMDPNLGQVQYVQFPQREDGIDRNDYANRNTVFF 598
Qy 616 DINLRGLDGGTQGVVYVGTGCVFNRTAIYGEPPIKA--KPGGLASLCCGGKKASKSKR 673
Db 673 DINLRGLDGGTQGVVYVGTGCVFNRTAIYGEPPIKA--KPGGLASLCCGGKKASKSKR 673
Db 599 DINLRGLDGGTQGVVYVGTGCVFNRTAIYGEPPIKA--KPGGLASLCCGGKKASKSKR 658
Qy 674 SSDKKKSNKHVDSVPFVNLIEDIEEGVAGFDEKSVLMSQMSLEKRGQSAFVASTL 733
Db 733 SSDKKKSNKHVDSVPFVNLIEDIEEGVAGFDEKSVLMSQMSLEKRGQSAFVASTL 733
Db 659 ESDKKKSGRHTDSTVPVFNLDIEEGVAGFDEKSVLMSQMSLEKRGQSAFVASTL 718
Qy 734 MEYGGVQSSPTPSLKEATHVISCYEDKSEMTGTEIGWYIGSVTEDIITGFKMHARGWR 793
Db 793 MEYGGVQSSPTPSLKEATHVISCYEDKSEMTGTEIGWYIGSVTEDIITGFKMHARGWR 793
Db 719 MENGVPSPATPENLLKEATHVISCYEDKSEMTGTEIGWYIGSVTEDIITGFKMHARGWR 778
Qy 794 SVYCMPLKPAFKGAPINLSDRLNQVLRWALGSVEILFSGHCPWLYGYGGRKFLERFAY 853
Db 853 SVYCMPLKPAFKGAPINLSDRLNQVLRWALGSVEILFSGHCPWLYGYGGRKFLERFAY 853
Db 779 SIYCMPLKPAFKGAPINLSDRLNQVLRWALGSVEILFSGHCPWLYGYGGRKFLERFAY 838
Qy 854 INTIYPLTSLPLLYVCLPAICLLGKFTMPISNLSIWFALFSLFATGILEMRWS 913
Db 913 INTIYPLTSLPLLYVCLPAICLLGKFTMPISNLSIWFALFSLFATGILEMRWS 913
Db 839 VNTIYPLTSLPLLYVCLPAICLLGKFTMPISNLSIWFALFSLFATGILEMRWS 898
Qy 914 GVGIDWNRNEQFVIGGISAHLFAVFGQLLKVLGADINFTVTSKANDDEGDFAEIYLF 973
Db 973 GVGIDWNRNEQFVIGGISAHLFAVFGQLLKVLGADINFTVTSKANDDEGDFAEIYLF 973
Db 899 GVGIDWNRNEQFVIGGISAHLFAVFGQLLKVLGADINFTVTSKANDDEGDFAEIYLF 958
Qy 974 KWTLLPPTTLLIINNVGVVAGTSVAINSGYOSWGLFGKLFFAFWIVHLYPFLKGLM 1033
Db 1033 KWTLLPPTTLLIINNVGVVAGTSVAINSGYOSWGLFGKLFFAFWIVHLYPFLKGLM 1033
Db 959 KWTLLPPTTLLIINNVGVVAGTSVAINSGYOSWGLFGKLFFAFWIVHLYPFLKGLM 1018
Qy 1034 GRQNRPTIIVVAVLLASTFSLWVRDPFTTRLAGPNQTCGINC 1080
Db 1080 GRQNRPTIIVVAVLLASTFSLWVRDPFTTRLAGPNQTCGINC 1080
Db 1019 GRQNRPTIIVVAVLLASTFSLWVRDPFTTRLAGPNQTCGINC 1065
```

RESULT 6

O48948

ID O48948 PRELIMINARY; PRT: 1065 AA.

```
AC O48948;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN ATH-B.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Eukaryota; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLOMBIA;
RX MEDLINE=98111412; PubMed=9445479;
RA Arioli T., Peng L., Betzner A.S., Burn J., Wittke W., Herth W.,
RA Camilleri C., Hofte H., Plazinski J., Birch R., Cork A., Glover J.,
RA Redmond J., Williamson R.E.;
RT "Molecular analysis of cellulose biosynthesis in Arabidopsis.";
RL Science 279:717-720(1998).
DR EMBL; AF027174; AAC39336.1; -.
DR InterPro; IPR001510; Cellulose_synt.
DR Pfam; PF03552; Cellulose_synt; 1.
DR SMART; SM00184; RING; 1.
DR DR
SQ SEQUENCE 1065 AA; 119659 MW; 1BF01C95FB85BE4C CRC64;
```

Query Match 78.0%; Score 4507; DB 10; Length 1065;  
Best Local Similarity 78.4%; Pred. No. 0;  
Matches 836; Conservative 98; Mismatches 107; Indels 26; Gaps 8;

```
Qy 19 CQICADGLTLDGDTACDVCRCPCYEHKEGTOACLOCKTKYKRRHSGPAIR 78
Db 20 CQICSDNVGKTVGDGRVACDICSFPVRCPCYEHKEGTOACLOCKTKYKRRHSGPAIR 79
Qy 79 GEGDDTDADDGS--DFNYPASGTEDQKQIADRMRSWRMTGSGNVGHPKYDSGEIGLS 137
Db 80 GKDDEGLADEGIVFNPY-----QKEKISERMLGWHLTRGKEEMGEPOYDK----- 127
Qy 138 KYDSEIPIRGVPSVTSNSQ--MSGEIPGASPDHMHMSTGNSIRAPFPY--VNHSPNPSRE 195
Db 128 -----EVSHNHLPLRTSRQDTSGEFSAASPERLSVSTTAGGRKRLYSDDVQSNRRIV 182
Qy 196 FSGSIVGNVANKERVGDGKMKQDKGAIPMTNGTSTIAPSEGRRAAIDIDASTEYNMEDALLND 255
Db 183 DPVGLGNVANKERVGDGKMKQDKGAIPMTNGTSTIAPSEGRRAAIDIDASTEYNMEDALLND 238
Qy 256 ETROPILSRKVPDIASSKINPYRMVIVLRLVLSIFLHYRLTNPNVRYNAYPLWLLSVICEIWF 315
Db 239 EAROPILSRKVPDIASSKINPYRMVIVLRLVLSIFLHYRLTNPNVRYNAYPLWLLSVICEIWF 298
Qy 316 ALSWILDOQFPKWPPIINRETYLDRLALRYDREGESQLAADVIFVSTVDPLKEPPIVTANT 375
Db 299 ALSWILDOQFPKWPPIINRETYLDRLALRYDREGESQLAADVIFVSTVDPLKEPPIVTANT 358
Qy 376 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWPFVFKYDIEPRAPFAYFC 435
Db 359 VLSILAVDYPVDKVCYVSDGASMLTFDALAETSEFARKWPFVFKYDIEPRAPFAYFC 418
Qy 436 QKIDYLDKDKVQPSFVKDORRAMKREYEFEKIRINALVSKALKVPEEGWIMODGTPWPGNNT 495
Db 419 AKIDYLDKDKVQPSFVKDORRAMKREYEFEKIRINALVSKALKVPEEGWIMODGTPWPGNNT 478
Qy 496 RDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKRPQFQHHKAGAMNALVRYSVLTNG 555
Db 479 GDHPGMIQVFLGHSGLDTEGNEPLRLVYVSREKRPQFQHHKAGAMNALVRYSVLTNG 538
Qy 556 QYMLNLCDDHYINNSKALREAMCFMDPNLGQVQYVQFPQREDGIDRNDYANRNTVFF 615
Db 539 PFILNLCDDHYINNSKALREAMCFMDPNLGQVQYVQFPQREDGIDRNDYANRNTVFF 598
Qy 616 DINLRGLDGGTQGVVYVGTGCVFNRTAIYGEPPIKA--KPGGLASLCCGGKKASKSKR 673
Db 673 DINLRGLDGGTQGVVYVGTGCVFNRTAIYGEPPIKA--KPGGLASLCCGGKKASKSKR 673
```

599	Db	DINLRGLDGIQGVYGTGCVFNRTALGYEPPIKVKHKKPSLLSKLGGSRKKNKAKK	658
674	QY	SSDKKSNKHVDSSVPVFNLEIDIEEGVEGAGFDDKESVLMSQMSLERFRFGQSAFAVASTL	733
659	Db	ESDKKSGRHTDSTVPVFNLDIDIEEGVEGAGFDDKALLMSQMSLERFRFGQSAFAVASTL	718
734	QY	MEYGGVQSSTPSLKEAHIIVSCGYEDKSEWGTETGWIYGSVTEIDILTGFKMHARGWR	793
719	Db	MENGVPSPATPENLLKEAHIIVSCGYEDKSDGMEIGWIYGSVTEIDILTGFKMHARGWR	778
794	QY	SVYCMRPAFAKSGAPINLSDRLNQVLWALGSVEILFSRHCPWYGYGRLKFLERFAY	853
779	Db	SIYCMPLPAFAKSGAPINLSDRLNQVLWALGSVEILFSRHCPWYGYGRLKFLERFAY	838
854	QY	INTTYPLTSLPLVYCYLPAICLLTGKFIIMPEISNLSIWFIALFSLFISFATGILEMRWS	913
839	Db	VNTTYPTISPLIMYCTLLAVCLFTNQFIIPQISNLSIWFISLFLSIFATGILEMRWS	898
914	QY	GVGIDEMWRNEQFVWIGGSAHLFAVFGQILKVLGIDTNTVTYSKANDDEGDAEALYMF	973
899	Db	GVGIDEMWRNEQFVWIGGSAHLFAVFGQILKVLGIDTNTVTYSKANDDEGDAEALYLF	958
974	QY	KWTTLLIPPTILLINMVGWVAGTSYAINSGYQSWGFLGKLFPAFWVIVHLIYFLFKGLM	1033
959	Db	KWTTLLIPPTILLINMVGWVAGTSYAINSGYQSWGFLGKLFPAFWVIVHLIYFLFKGLM	1018
1034	QY	GRQNRPTTIVWAVLLASIFSLWVRDPTTBLAGPNTOTCGINC	1080
1019	Db	GRQNRPTTIVWVSLLASIFSLWVRIDPFTTBLAGPNTOTCGINC	1065
RESULT 7			
Q9LLI8		PRELIMINARY; PRT; 1074 AA.	
ID	Q9LLI8;		
AC	Q9LLI8;		
DT	01-OCT-2000 (TrEMBLrel. 15, Created)		
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)		
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)		
DE	Cellulose synthase-2.		
GN	CESA-2.		
OS	Zea mays (Maize).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;		
OC	Panicoidae; Andropogoneae; Zea.		
OX	NCBI_TaxID=4577;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE-20398328; PubMed=10938350;		
RA	Holland N., Holland D., Helentjaris T., Dhugga K.S.,		
RA	Xoconostle-Cazares B., Delmer D.P.;		
RT	"A comparative analysis of the plant cellulose synthase (Cesa) gene		
RT	family";		
RL	Plant Physiol. 123:1313-1324(2000).		
DR	EWBL: AF200526; AAF89962.1; "		
DR	InterPro: IPR005150; Cellulose_synt.		
DR	InterPro: IPR001841; Znf_fing.		
DR	Pfam: PF03552; Cellulose_synt; 1.		
DR	SMART: SM00184; RING; 1.		
SQ	SEQUENCE 1074 AA; 121187 MW; A116E66F0564E210 CRC64;		
Query Match			
Best Local Similarity 69.1%; Score 3993.5; DB 10; Length 1074;			
Matches 749; Conservative 134; Mismatches 153; Indels 59; Gaps 20;			
QY	2	DGDADA-LKSGRHAGDVCQICADGLTTLDDGVFTACDVCRRFCYERHERKEGTQA	60
Db	23	DGDAPVPAKPTKSANGQVCICGTGTGVSATGDGVFACNECAFPVCRPCYERKEGNQC	82
QY	61	CLQCKTKYKRRGSPALREGEEDTDADD-GSDENYASGTEDQKQIADRMRSWRMTG	119
Db	83	CPOCKTRYKKGKSPRVHGGD-EEDVDLDNEFN-----KQ-----	119
QY	120	GSGNVGHPKY----DSGEIGLS-KYDSGEIPRGVPSVTN-SQMSGEIPGASPDHH-MMS	172

Db	120	--GNGKPEWLOQDDADLSS	173
Qy	173	PTGN-ISRRAFPVFNHSPNS	230
Db	174	PTSSYDPSVPV- -RIVDS	226
Qy	231	PSEGRAATDADSTEYNKDAL	290
Db	227	-KYPEARGDME-GTGSNGEDM	284
Qy	291	HYRLTNVRNAYPLWLLSVIC	350
Db	285	QYRISHPRNAYGLWVSVIC	344
Qy	351	QLAAVDIFSVTDPLKEPPIT	410
Db	345	QLAPIDVFVSTVDPLKEPLIT	404
Qy	411	EFARKWTFPVKKYKIDIEPAP	470
Db	405	EFARKWTFPVCKKHIEPAP	464
Qy	471	VSKALKVPEBGWINQDTPWG	530
Db	465	VAKAQVPEBGWIMADGTWPG	524
Qy	531	PGFOHKKAGAMNALRVSAV	590
Db	525	PGFOHKKAGAMNALRVSAV	584
Qy	591	YVQFPQRFQDIDRNDRYANR	650
Db	585	YVQFPQRFQDIDLHDYANR	644
Qy	651	AK--KPGFLASLGGCKKASK	708
Db	645	EADLEPNIVKSCGRK-RKN	701
Qy	709	KSVLMSQMSLEKREGQSAFVA	768
Db	702	RSVLMSRKLEKREGQSPFIA	761
Qy	769	EIGHYGSVTEIDILTGFKMH	828
Db	762	EIGHYGSVTEIDILTGFKMH	821
Qy	829	ILFSRHCPWYGYGRLKFLER	888
Db	822	ILFSRHCPWYGYGRLKFLER	881
Qy	889	NLASIFIALFSLFATGILE	948
Db	882	NYAGMEFILLFASIFATGILE	941
Qy	949	GIDTNFTVTSKANDDEGDAE	1008
Db	942	GIDTNFTVTSKANDDEGDAE	1001
Qy	1009	GPLFGKLFPAFWVIVHLIYFL	1065
Db	1002	GPLFGKLFPAFWVIVHLIYFL	1061
Qy	1066	TRLAGPNIQTGINC	1080
Db	1062	QKAAA--LGQGVNC	1074

RESULT 8  
Q9LLI9  
ID Q9LLI9  
AC Q9LLI9;  
DT 01-OCT-2000 (TrEMBLrel. 15  
PRT; 1075 AA.  
ated)





RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF027172; AAC39334.1; -  
DR EMBL; AL034567; CAA22568.1; -  
DR EMBL; AL161581; CAB79958.1; -  
DR InterPro; IPR001841; Znf\_ring.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF03552; Cellulose\_synt; 1.  
DR SMART; SM00184; RING; 1.  
SQ SEQUENCE 1081 AA; 122236 MW; BDEB5D9DEB334D59 CRC64;  
  
Query Match 68.6%; Score 3963.5; DB 10; Length 1081;  
Best Local Similarity 68.0%; Pred. No. 3e-312;  
Matches 733; Conservative 138; Mismatches 168; Indels 39; Gaps 15;  
  
QY 2 DGADALSGRHGAGDVQCICADGLGTTLDGVDVFTACDVCRFPVCRPCYERHERKEGTQAC 61  
DB 25 DGGTKPLKNN--GQICQCGDDVGLAETGDFVACNECAFPVCRPCYERHERKDGTCQC 81  
  
QY 62 LOCKTKYKRHRGSPAIRGEEDDTADGSDENYPASGTEDOKOKIADMRSMRMTGGS 121  
DB 82 POCKTRFRHRGSPRVEGDEDDVDVDFNEFNAQAGANKRHRGHEEFSS----- 133  
  
QY 122 GNVGHPKXDSGEIGLSKYD---SCEIPRGVVPVTSQMSGEIPGASPDHMMSPGTNIS 178  
DB 134 ---SSRHESQIPLLTGHTVSGEIRTPDQSVRTT--SGPL-GPSDRNAISSP--YID 184  
  
QY 179 RRAFPVYVNIHSPNDRF--SGSGTGNVANKERVDGKWKQKGAIPWTNGTISIAPSEGRAA 237  
DB 185 PRQVPV--RIVDPKDLNYSGLGNDVKERVEGKLGKQKNLQMTG----KYHEGKG 238  
  
QY 238 TDIDASTEYNNEDALLNDETQPLSRKVPVTSKINPRVIVLRLVVLGFIHYRTNP 297  
DB 239 -EIE-GTSGNGEELQADDTLPMRSVVPVPSRLTYPVVVILRLIILCFQYRTHP 296  
  
QY 298 VRNAYPLWLSVCEIWFALSWILDQPKWFPNINRETYLDRLALRYDREGESQLAADI 357  
DB 297 VKNAYPLWLSVCEIWFALSWILDQPKWFPNINRETYLDRLALRYDREGESQLPVVD 356  
  
QY 358 FVSTVDLPKPPVITANTVLSILADVDPDKVSCYVSDDGASMLTFDALAETSEFAKRW 417  
DB 357 FVSTVDLPKPPVITANTVLSILADVDPDKVSCYVSDDGASMLTFDALAETSEFAKRW 416  
  
QY 418 PFVKKYDIEPRAPFEYFCQKIDYLDKQVPSFKVDRAMKREYEEFKIRINALYSKALKY 477  
DB 417 PFCKKFIIEPRAPFEYFAQKIDYLDKQVPSFKVDRAMKREYEEFKIRINALYSKALKY 476  
  
QY 478 PEEGWIMODGTPPNNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGFOHK 537  
DB 477 PEEGWIMODGTPPNNTRDHPGMIQVFLGHSGLDTEGNEPLRLVYVSRKRPFGFOHK 536  
  
QY 538 KAGANMALRVSAVLITNGQYMLNDCDHYITNNSKAVREACFMMDPNLGPQVQVQPQR 597  
DB 537 KAGANMALRVSAVLITNGAYLLNVCDHYENNSKAKEACFMMDPAIGKKCCYVQPQR 596  
  
QY 598 FDGIDRNDYANRNTVFNDRLDGIQGVVYVGTGCVENRTAIYEPPIKAK--KPG 655  
DB 597 FDGIDRNDYANRNTVFNDRLDGIQGVVYVGTGCVENRTAIYEPPIKAK--KPG 656  
  
QY 656 FLA-SLGGKKKASKKSSDKSKNKHVDVSVVPLENLEDIEGVBGAGFDDEKSVLMS 714  
DB 657 IIVKSCGSKKSKSKSKYKRYKRNLSRSDSNAPLNFMEDIDEGFE--GYDDERSLMS 714  
  
QY 715 QMSLEKRFQGSAAVSTLMYEGVQSPSTPESLLKBAIHVSCGYEDKSEWTEIGWIY 774  
DB 715 QRSVEKRFQGSAAVSTLMYEGVQSPSTPESLLKBAIHVSCGYEDKSEWTEIGWIY 774  
  
QY 775 GSVTEDILTGFKHARGWRSYCMKRPAPKPGAPINLSRLNQLRWALGSEIILSRH 834  
DB 775 GSVTEDILTGFKHARGWRSYCMKRPAPKPGAPINLSRLNQLRWALGSEIILSRH 834  
  
QY 835 CPLWYGGRLKFLERFAYINTTITPLTSLVYICLPAICLLTGKFINPEISNLASIW 894

DB 835 CPIWYGYHGRLLRLERAYINT 894  
QY 895 FTALFLSIFATGILEMRWQVGV 954  
DB 895 FILLFISIAVTGILELRWQVGS 954  
QY 955 TVTISKANDEGDFAEILYMKWT 1014  
DB 955 TVTISKATDEGDFAEILYMKWT 1014  
QY 1015 LFFAFVVIHLYPFLKGLMGRO 1072  
DB 1015 LFFALVIAHLYPFLKGLMGRO 1070  
  
RESULT 10  
Q9LLI2 PRELIMINARY;  
AC Q9LLI2; ST; 1094 AA.  
DT 01-OCT-2000 (TrEMBLrel. 15; ted)  
DT 01-OCT-2000 (TrEMBLrel. 15; sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21; annotation update)  
DE Cellulose synthase-8.  
GN CESA-8.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; S zophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyt liopsida; Poales; Poaceae; PACC clade;  
OC Panicoidae; Andropogoneae?  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20398328; PubMed=10 50;  
RA Holland N., Holland D., Hel aris T., Dhugga K.S.,  
RA Xoonosle-Cazares B., Delm P.;  
RT "A comparative analysis of olant cellulose synthase (Cesa) gene family.";  
RL Plant Physiol. 123:1313-132 00).  
DR EMBL; AF200532; AAF89968.1; synt.  
DR InterPro; IPR001841; Znf\_r  
DR Pfam; PF03552; Cellulose\_sy  
DR SMART; SM00184; RING; 1.  
SQ SEQUENCE 1094 AA; 122575 DFBBC18F49D23F5E CRC64;  
  
Query Match 66.5%;  
Best Local Similarity 65.4%;  
Matches 723; Conservative 15;  
  
QY 9 KSRHGHAGDV-----CQICAE 62  
DB 25 ESGAGGGAARRAPQCICGE 84  
  
QY 63 QCKTKYKRHRGSPAIRGEED 121  
DB 85 QCRTYKRLKCCRVAGDEEE 144  
  
QY 122 GNVGHPKXDSGEIGLSKYDSE 175  
DB 145 GD-AHP-----GI 186  
  
QY 176 NISRRAPFPVYN-----HS 227  
DB 187 GGRHPLPPADNLPVQPRSE 238  
  
QY 228 SIAPSEGRATDIDASTEYNN 287  
DB 239 QHVRSEGGDWDGDA----- 292  
  
QY 288 IFHYRLTNVNRNAYPLWLS 347  
DB 293 FFFHYRVMHPAKDAFALWLS 352

QY 348 EPSQLAAVDIFVSTVDLPKEPIVNTANTVLSILAVDYPVDKVCYVSDDGASMLTFDALA 407  
Db 353 QFSQLAPIDFFVSTVDLPKEPIVNTANTVLSILSVDEYKVCYVSDDGAAMLTFEALS 412  
QY 408 ETSEFARKWVPFKKYDIEPRAPEYFQKIDYLDKDKYQPSFVKDRRAKREYEEFKIRI 467  
Db 413 ETSEFARKWVPFKKYDIEPRAPEYFQKIDYLDKDKVAASFVRERRAKREYEEFKIRI 472  
QY 468 NALVSKALVPRGGWIMQDGTWPGNNTRDHPGMIOVFLGHSGGLDTGEGNELPLVTVSR 527  
Db 473 NALVARAKVPEGWIMQDGTWPGNNTRDHPGMIOVFLGHSGGLDTGEGNELPLVTVSR 532  
QY 528 EKRPQGHKKKAGAMNALVRVSAVLNGQYMLNLDCHYINNSKAVREAMCFMLDPNLG 587  
Db 533 EKRPQGHKKKAGAMNALVRVSAVLNAAAYLLNLDCHYINNSKAIKAMCFMMDPLVGK 592  
QY 588 QVCYVQFQFQDIDRNDYARNVTFEDINLRGLDQGVVYVGTCTCFVNRTAIYGP 647  
Db 593 KVCYVQFQFQDIDRNDYARNVTFEDINLRGLDQGVVYVGTCTCFVNRQALYGYDA 652  
QY 648 PIKAKRP-----GFLASLCGGKKKASKKRSDDKKKS--NKHVDSVVPVFNLED 695  
Db 653 P-KTKKPPSKTCNCPKWCLSCCSRNKNKKTTPKTEKKKLFKKAENPSPAYALGE 711  
QY 696 IEBGVGAGFDDKSVLMSQMSLEKRFQGSAAFAVASTLMKEYGGVQPSSTPESILKEAIV 755  
Db 712 IDEGAPGA--DIEKAGIVNQKLEKFKQSSVFAVASTLLENGGTLKSPASLKEAIV 769  
QY 756 ISGCGYEDKSEWTEIGWYGSVTEIDILTGFKMHARGWSVYCMKPRPAFKGSAPINLSDR 815  
Db 770 ISGCGYEDKDWCKEIGWYGSVTEIDILTGFKMHARGWSVYCMKPRPAFKGSAPINLSDR 829  
QY 816 LNOVLRLWALSGVEILFHSRCHPLWYGGYGRKFLERFAINTIITPLSLPLLVTCILPAI 875  
Db 830 LHOVLRLWALSGVEILFHSRCHPLWYGGYGRKFLERFINSIVPMTSIPLLAYCTIPAI 889  
QY 876 CLLTGXFIMPEISNLASLWIFALFLSIFATGILEMRWSGVGIDEWNRNEQFWIGGISAH 935  
Db 890 CLLTGXFITPELTNVASLWIFALFLSIFATGILEMRWSGVGIDEWNRNEQFWIGGISAH 949  
QY 936 LFAVFGLLKVLAGIDTNETVNSKANDEGDFAEALYMKFWTLLIPPTTILIIINWGVVA 995  
Db 950 LFAVFGLLKVLAGIDTNETVNSKANDEGDFAEALYMKFWTLLIPPTTILIIINWGVVA 1008  
QY 996 GTSYAINSGYQSGWPLFGKLFPAFWIVHLXPLFKGLMGRONRPTIIVWAVLLASIFS 1055  
Db 1009 GISNALNGYESWGLPFGKLFPAFWIVHLXPLFKGLMGRONRPTIIVWAVLLASIFS 1068  
QY 1056 LLWVRVDPETTRLAGNIQTCCINC 1080  
Db 1069 LLWVRVDPFLAKSNGPLLEECGLDC 1093

## RESULT 11

Q9LLI4  
ID Q9LLI4 PRELIMINARY; PRT: 1059 AA.  
AC Q9LLI4:  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Cellulose synthase-6.  
GN CESA-6.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20398328; PubMed=10938350;  
RA Holland N., Holland D., Helentjaris T., Dhugga K.S.,  
RA Xoonostle-Cazares B., Delmer D.P.;  
RT "A comparative analysis of the plant cellulose synthase (CesA) gene

RT family";  
RL Plant Physiol. 123:1313-1324 (2000).  
DR EMBL: AF200530; AAF99966.1; -.  
DR InterPro: IPR00150; Cellulose\_synt.  
DR InterPro: IPR001841; Znf\_ring.  
DR Pfam: PF03552; Cellulose\_synt; 1.  
DR SMART: SM00184; RING; 1.  
SQ SEQUENCE 1059 AA; 120044 MW; 289DA26B25232249 CRC64;  
  
Query Match 66.1%; Score 3820; DB 10; Length 1059;  
Best Local Similarity 65.1%; Pred. No. 1.3e-300;  
Matches 714; Conservative 140; Mismatches 168; Indels 74; Gaps 16;  
  
QY 16 GDVQCICADGLGTTLDGVFTACDVCREFPVCRCPCYEHKKEGTQACLOCKTKYKRHRGSP 75  
Db 6 GQVCQICGDDVGRNPDGEPFVACNECAPICDCYCYEYERREGTQNCPOCKTRFRKFGCA 65  
QY 76 ATRGEGDDTDADGSDRNYPASGTEDQOKIADRMRSWRMNTGSGNVG-----HPKY 129  
Db 66 RVPGDEEDGVDDLENEFNW---SDKHDSQYLAESMLHAHMSYGRGADLDGVPPQFHP-- 120  
QY 130 DSGEIGLSKYDSGEIPRGVYPSVTNSQSGEIPGASPDHMMSP---GNISRRAPPYV 186  
Db 121 -----IPN--VPLLTNGQWDDIP---PQHALVPSFVGGGGRKRLHPLPYA 161  
QY 187 N-----HSPNPSREFSG-SIGNVAKRERVDGWMKMDKGAIPMTNGTSTAPSEGRAAT 238  
Db 162 DPNLPPQPRSDPDKLAAYGYGVAMKERMESWKQER--MHQTR-----NDGGDD 213  
QY 239 DIDASTEYNMEDALNDETROPLSRKVPTASSKINPYBMVIVLRLVLSIFLHYRLNPV 298  
Db 214 GDAA-----DLPLMEARQPLSRKIPLPSSQINPYRMIIILRLVLVLCFFHVRVHPV 266  
QY 299 RNAYPLWLLSVTCEIWFALSWILDQPKWFPINRETYLDRLLARYDREGESQLAADVIF 358  
Db 267 PDFAFWLLSVTCEIWFAMSWILDQPKWFPITERETYLDRLSLRFDEKBPQLAPVDF 326  
QY 359 VSTVDPLKPPVTANTVLSILAVDYPVDKVCYVSDDGASMLTFDALAESEFARKWVP 418  
Db 327 VSTVDPLKPPVTANTVLSILSVDPYVDKVCYVSDDGAAAMLFEALSESEFAKKWVP 386  
QY 419 FVKYDIEPRAPEYFQKIDYLDKDKVQPSFVKDRRAKREYEEFKIRINALVSKALKVP 478  
Db 387 FCKRYSLEPRAPEYFQKIDYLDKDKVAPNFEVRRERAKREYEEFKIRINALVAKAKVP 446  
QY 479 EGEWIMQDGTWPGNNTRDHPGMIOVFLGHSGGLDTGEGNELPLRVYVYSRERKPGFHHKK 538  
Db 447 EGEWIMQDGTWPGNNTRDHPGMIOVFLGHSGGLDTGEGNELPLRVYVYSRERKPGFHHKK 506  
QY 539 AGAMNALVRVSAVLNGQYMLNLDCHYINNSKAVREAMCFMLDPNLGQVYVQFQF 598  
Db 507 AGAMNALVRVSAVLNAPYLLNLDCHYINNSKAIKAMCFMMDPLLGKVKYVQFQF 566  
QY 599 DGIIDRNDYARNVTFEDINLRGLDQGVVYVGTCTCFVNRTAIYGPPTTAKKP--- 654  
Db 567 DGIIDRNDYARNVTFEDINLRGLDQGVVYVGTCTCFVNRQALYGYDAP-KTKKPSRT 625  
QY 655 -----GFLASLCGGKKKASKKRSDDK-----KKSNNHVDSSVVPVFNLEDIEEGVAG 704  
Db 626 CNCWPKWICCCCFGNRRKTKTKTKSKPKFEKIKKFKKKNQAPAYALGEAAPGA- 684  
QY 705 FDDEKSVLMSQMSLEKRFQGSAAFAVASTLMKEYGGVQPSSTPESILKEAIVISGYEDKS 764  
Db 685 -ENEKASIVNQKLEKFKQSSVFAVASTLLENGGTLKSPASLKEAIVISGYEDKT 743  
QY 765 EWGTEIGWYGSVTEIDILTGFKMHARGWSVYCMKPRPAFKGSAPINLSRLNQLRWAL 824  
Db 744 GWKDGIGWYGSVTEIDILTGFKMHARGWSVYCMKPRPAFKGSAPINLSRDFHQLRWAL 803  
QY 825 GSVEILFSPCHPLWYGGYGRKFLERFAINTIITPLSLPLLVTCILPAICLTGKFEIM 884  
Db 804 GSVEILFSPCHPLWYGGYGRKFLERFAINTIITPLSLPLLVTCILPAICLTGKFEIM 863



RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose  
RT synthase required for secondary cell wall synthesis.";  
RL Plant Cell 11:769-780(1999).

```

RN [2]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL [3]
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF091713; AAD40885.1; -
DR EMBL: AL391142; CAC01737.1; -
DR InterPro: IPR005150; Cellulose_synt.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synt; 1.
DR SMART: SM00184; RING; 1.
SQ SEQUENCE 1026 AA; 115797 MW; 503BFBC78BB6E511 CRC64;

Query Match 55.8%; Score 3801; DB 10; Length 1026;
Best Local Similarity 67.5%; Pred. No. 4.2e-299;
Matches 725; Conservative 104; Mismatches 156; Indels 88; Gaps 17;

QY 16 GDVCOICADGLGTTLDGDFVTCACDVCRCPCYEHKEGEGTQACLOCKTKYKRRGSP 75
DB 16 GDVCOICADGLGTTLDGDFVTCACDVCRCPCYEHKEGEGTQACLOCKTKYKRRGSP 75
QY 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPOCKTRKRLRGSP 93
DB 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPOCKTRKRLRGSP 93
QY 76 AIRGEGDDTDADGSDFNYPASGTEDQOKIADMRSMRMNTGSGNGHPKYDSGEIG 135
DB 76 AIRGEGDDTDADGSDFNYPASGTEDQOKIADMRSMRMNTGSGNGHPKYDSGEIG 135
QY 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142
DB 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142
QY 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRRAPFPVYHSPNP 192
DB 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRRAPFPVYHSPNP 192
QY 143 -----NGRFPVPVIAGHSGEFPVGGYGNGEH-----GLHKRV-----HPYP 179
DB 143 -----NGRFPVPVIAGHSGEFPVGGYGNGEH-----GLHKRV-----HPYP 179
QY 193 SRFSGSIGNAVKEKRVGDKMKDKGAIPMTNGTSIAPSEGRAADIDIDASTEYNMEDAL 252
DB 193 SRFSGSIGNAVKEKRVGDKMKDKGAIPMTNGTSIAPSEGRAADIDIDASTEYNMEDAL 252
QY 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG-----214
DB 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG-----214
QY 253 LNDETROPLSRKVPDIASSKINPVMIVLRLVLSIFLHYRLTNPVNNAIPLWLSVICE 312
DB 253 LNDETROPLSRKVPDIASSKINPVMIVLRLVLSIFLHYRLTNPVNNAIPLWLSVICE 312
QY 215 LIDEAROLSRKVPDIASSKINPVMIVLRLVLSIFLHYRLTNPVNNAIPLWLSVICE 274
DB 215 LIDEAROLSRKVPDIASSKINPVMIVLRLVLSIFLHYRLTNPVNNAIPLWLSVICE 274
QY 313 IWFALSILDOFPKWPPIERETYLRLSLRYEREGERPNMLAPVDVSVSDPLKEPIYV 372
DB 313 IWFALSILDOFPKWPPIERETYLRLSLRYEREGERPNMLAPVDVSVSDPLKEPIYV 372
QY 275 IWFALSILDOFPKWPPIERETYLRLSLRYEREGERPNMLAPVDVSVSDPLKEPIYV 334
DB 275 IWFALSILDOFPKWPPIERETYLRLSLRYEREGERPNMLAPVDVSVSDPLKEPIYV 334
QY 373 ANTVLSILAVDYPVDKVCVSDGASMLTFDALAEISFARKWPVFKKYDIEPRAPEF 432
DB 373 ANTVLSILAVDYPVDKVCVSDGASMLTFDALAEISFARKWPVFKKYDIEPRAPEF 432
QY 335 SNTVLSILAMDYPVEKISCVSDGASMLTFESLSETAFARKWPVFKKFSSTEPRAPEM 394
DB 335 SNTVLSILAMDYPVEKISCVSDGASMLTFESLSETAFARKWPVFKKFSSTEPRAPEM 394
QY 433 YFCOKIDYLDKQVPSVKDRRAKREYEFEKIRINALYSKALKVPPEGWIMODGTPWG 492
DB 433 YFCOKIDYLDKQVPSVKDRRAKREYEFEKIRINALYSKALKVPPEGWIMODGTPWG 492
QY 395 YFLKVDYLDQKVPVFKERRAKREYEFEKIRINAKVASKVPLEGIMODGTPWG 454
DB 395 YFLKVDYLDQKVPVFKERRAKREYEFEKIRINAKVASKVPLEGIMODGTPWG 454
QY 493 NNTRDHPGMIQVFLHSGGGLDTGNEPLRLVYVSREKRPFGQHHKAGANNAIYVRSVYL 552
DB 493 NNTRDHPGMIQVFLHSGGGLDTGNEPLRLVYVSREKRPFGQHHKAGANNAIYVRSVYL 552
QY 455 NNTKDHPGMIQVFLHSGGGLDTGNEPLRLVYVSREKRPFGQHHKAGANNAIYVRSVYL 514
DB 455 NNTKDHPGMIQVFLHSGGGLDTGNEPLRLVYVSREKRPFGQHHKAGANNAIYVRSVYL 514
QY 553 TNGOYMLNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 612
DB 553 TNGOYMLNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 612
QY 515 TNAPFMNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 574
DB 515 TNAPFMNLCDDHYNNKAVRAMCFMDPNLGPQVQVYQFPQDFGIDRNDRYANRNT 574
QY 613 VFEDINLRGLDGTQGPVYVGTGVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 668
DB 613 VFEDINLRGLDGTQGPVYVGTGVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 668
QY 575 VFEDINMKGLDGTQGPVYVGTGVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 632
DB 575 VFEDINMKGLDGTQGPVYVGTGVFNRTAIYGYEPPKAKKPGFLASLC-----GKKKAS 632
QY 669 KSKRRSDKKKSHVDSSVPVFNLEDIEEGVEGAGFDEKSVMSQMSLEKRFQGSAAF 728
DB 669 KSKRRSDKKKSHVDSSVPVFNLEDIEEGVEGAGFDEKSVMSQMSLEKRFQGSAAF 728
QY 633 -----NKKFSKNDMNGDVAALG-----GABG-----DKEHLMSEMFEXTFGQSSIF 674
DB 633 -----NKKFSKNDMNGDVAALG-----GABG-----DKEHLMSEMFEXTFGQSSIF 674
QY 729 VASTLMEYGVGPOSSPESILLKEAIHIVSCGYEDKSEWGTETGWIYGSVTEDILTGFKMH 788
DB 729 VASTLMEYGVGPOSSPESILLKEAIHIVSCGYEDKSEWGTETGWIYGSVTEDILTGFKMH 788
```

```

DB 675 VTSTLMEEGVPPSSSPAVLLKEAIHIVSCGYEDKTEGTGELGIYGSITEDILTGFKMH 734
QY 789 ARGWRSVYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWGYG--GGRKLF 847
DB 735 CRGWSIYCMKRPAPKGSAPINLSDRNLQVLRWALGSVEILFSRHCHPLWGYGKGKLV 794
QY 848 LERFAYINTTYPTSLPLLYCICLPAICLLTGKFIPEISNLASIFALFISFATGI 907
DB 795 LERFAYANTTYPTSLPLLYCICLPAICLLTGKFIPEISNLASIFALFISFATGI 854
QY 908 LEMRWSGVGIDEMWRNEQFWIGGISAHLFAVFOGLLKLVLGIDTNTVTSKANDESGDF 967
DB 855 LELRWSGVGIDEMWRNEQFWIGGISAHLFAVFOGLLKLVLGIDTNTVTSKANDESGDF 913
QY 968 AELYMKFTWTLTLLIPPTTILINMVGWVAGTSYAINSGYSQSWGPLFGKLFPAFWIVHLYP 1027
DB 914 GELYAFKWTLLIPPTTILINMVGWVAGTSYAINSGYSQSWGPLFGKLFPAFWIVHLYP 973
QY 1028 FLKGLMGQRNRTPTIIVYVAVLLASIFSLWVRVDPFTTTRLAGNPQTGCGINC 1080
DB 974 FLKGLMGQRNRTPTIIVYVAVLLASIFSLWVRVDPFTTTRLAGNPQTGCGINC 1026

RESULT 14
QYXHP6 PRELIMINARY; PRT; 1026 AA.
AC QXHP6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cellulose synthase catalytic subunit.
GN IRX3.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsals.
ON NCBI_taxID=3702;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. LANDSBERG RECTA;
RX MEDLINE=99264300; PubMed=10330464;
RA Taylor N.G., Scheible W.R., Cutler S., Somerville C.R., Turner S.R.;
RT "The irregular xylem3 locus of Arabidopsis encodes a cellulose
RL synthase required for secondary cell wall synthesis.";
DB EMBL: AF088917; AAD32031.1; -
DR InterPro: IPR005150; Cellulose_synt.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF03552; Cellulose_synt; 1.
DR SMART: SM00184; RING; 1.
SQ SEQUENCE 1026 AA; 115858 MW; 453BFD1D283CD4D70 CRC64;
```

Query Match 65.7%; Score 3795; DB 10; Length 1026;  
Best Local Similarity 67.5%; Pred. No. 1.3e-298;  
Matches 724; Conservative 104; Mismatches 157; Indels 88; Gaps 17;

```

QY 16 GDVCOICADGLGTTLDGDFVTCACDVCRCPCYEHKEGEGTQACLOCKTKYKRRGSP 75
DB 16 GDVCOICADGLGTTLDGDFVTCACDVCRCPCYEHKEGEGTQACLOCKTKYKRRGSP 75
QY 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPOCKTRKRLRGSP 93
DB 34 GQFCICGDIQLTVEGDLFVACNEGFPACRCYERREGTQNCPOCKTRKRLRGSP 93
QY 76 AIRGEGDDTDADGSDFNYPASGTEDQOKIADMRSMRMNTGSGNGHPKYDSGEIG 135
DB 76 AIRGEGDDTDADGSDFNYPASGTEDQOKIADMRSMRMNTGSGNGHPKYDSGEIG 135
QY 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142
DB 94 RVEGDE-DEEDIDD-IEYEFNIEHQDKHKSAAEAMLYGKMSYGRG-----PEDDE- 142
QY 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRRAPFPVYHSPNP 192
DB 136 LSKYDSGEIPRGVPSVTNSOMSGEIP--GASPDHMMSPGNISSRRAPFPVYHSPNP 192
QY 143 -----NGRFPVPVIAGHSGEFPVGGYGNGEH-----GLHKRV-----HPYP 179
DB 143 -----NGRFPVPVIAGHSGEFPVGGYGNGEH-----GLHKRV-----HPYP 179
QY 193 SRFSGSIGNAVKEKRVGDKMKDKGAIPMTNGTSIAPSEGRAATDIDASTEYNMEDAL 252
DB 193 SRFSGSIGNAVKEKRVGDKMKDKGAIPMTNGTSIAPSEGRAATDIDASTEYNMEDAL 252
QY 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG-----214
DB 180 SSE-AGSEG--GWRERDDWKLG-----HGNLGPEDDDPEMG-----214
```

QY 253 LNDETROPLSRKVPPIASSKINPYRMVIVLRLVLSIFLHYRLNPNVRNAYPLMLLSVICE 312  
Db 215 LIDEARQPLSRKVPPIASSKINPYRMVIVLRLVLSIFLHYRLNPNVRNAYPLMLLSVICE 274  
QY 313 IFALSWILDQFPKWPPIINRETYDLRLALRYDREGPSQALAAVDIFVSTVDPLKEPIVT 372  
Db 275 IFALSWILDQFPKWPPIINRETYDLRLALRYDREGPSQALAAVDIFVSTVDPLKEPIVT 334  
QY 373 ANTVLSILAVDYPVDKVCVSDGASMLTFDALAETSEFARKWVPVKYDYTEPRAPEF 432  
Db 335 SNTVLSILAVDYPVDKVCVSDGASMLTFDALAETSEFARKWVPVKYDYTEPRAPEM 394  
QY 433 YFCQKIDYLDKQVPQFVDRRRAMKREYEEFKIRINALVSKALKVPEEGWIMQDGPWP 492  
Db 395 YFTLKVDYLDQKHVPTFVKERRAMKREYEEFKIRINALVSKALKVPEEGWIMQDGPWP 454  
QY 493 NNTDRHPGMIQVFLHSGGLDTSGNELPRLVYVSRKRGFGHKKAGAMNALVRVAVL 552  
Db 455 NNTDRHPGMIQVFLHSGGLDTSGNELPRLVYVSRKRGFGHKKAGAMNALVRVAVL 514  
QY 553 TNGQYMLNLDCHYINNSKAVREAMCFMDPNLGPQVYVQFQFQFQFQFQFQFQFQFQF 612  
Db 515 TNAPFMLNLDCHYINNSKAVREAMCFMDPNLGPQVYVQFQFQFQFQFQFQFQFQFQF 574  
QY 613 VFEDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPPKAKKPGFLASLC----GGKKAS 668  
Db 575 VFEDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPPKAKKPGFLASLC----GGKKAS 632  
QY 669 KSKRSDDKKSNKHVDSSVPPNLEDEIEGVEGAGFDDKSVLMSOMSKLEKRFQSAAF 728  
Db 633 -----NKKFSKNDMGDVAALG-----GAEQ-----DKEHLMFMENFEKTFGOSSIF 674  
QY 729 VASTLMEYGVGPOSSPESILLKAIHVISCGYEDKESWGTGTEIGWYGSVTEIDILTGFKM 788  
Db 675 VTSLMEYGVGPOSSPESILLKAIHVISCGYEDKESWGTGTEIGWYGSVTEIDILTGFKM 734  
QY 789 ARGWRSYCMKPAKPAKGSAPINLSDRNLQVLRWALGVSVEILSRHCPLWYV--GGRLKF 847  
Db 735 CRGWSYCMKPAKPAKGSAPINLSDRNLQVLRWALGVSVEILSRHCPLWYV--GGRLKF 794  
QY 848 LERFAYINTIYPTSLPLVAYCILPAICLLTGKFTMPETSNLASIWFALFSLIFATGI 907  
Db 795 LERFAYINTIYPTSLPLVAYCILPAICLLTGKFTMPETSNLASIWFALFSLIFATGI 854  
QY 908 LEMRWGSGIDEWWRNEQFWVIGGISAHFLFAVFGGLKVLGAGIDTNTFTVTSKANDERGF 967  
Db 855 LEMRWGSGIDEWWRNEQFWVIGGISAHFLFAVFGGLKVLGAGIDTNTFTVTSKANDERGF 913  
QY 968 AELYMFKWTTLLPPTTILINMVGWVAGTSYAINSGYQSWGDLFGKLFATFVAVVHLYP 1027  
Db 914 GELYAFKWTLLPPTTILINMVGWVAGTSYAINSGYQSWGDLFGKLFATFVAVVHLYP 973  
QY 1028 FLKGLMGRQNTPTIVVAVLASIFSLWVRVDPPTTLRAGFNIGTCGINC 1080  
Db 974 FLKGLMGRQNTPTIVVAVLASIFSLWVRVDPPTTLRAGFNIGTCGINC 1026  
RESULT 15  
Q9SKJ5 PRELIMINARY; PRT; 1065 AA.  
AC Q9SKJ5;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE Putative cellulose synthase catalytic subunit.  
GN AT2G25540.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=20083487; PubMed=10  
RA Lin X., Kaul S., Rounsley S.  
RA Fujii C.Y., Mason T.M., Bow  
RA Buell C.R., Ketchum K.A., D  
RA Cronin L.A., Shen M., Van  
RA Adams M.D., Carrera A.J., G  
RA Copenhaver G.P., Preuss D.  
RA Salzberg S.L., Fraser C.M.,  
RT "Sequence and analysis of C  
RT thaliana.";  
RL Nature 402:761-768(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X.;  
RL Submitted (MAR-2000) to the  
DR EMBL; AC006300; AAD20713.1;  
DR InterPro; IPR005150; Cellul  
DR InterPro; IPR001309; ICE\_P2  
DR InterPro; IPR001841; Znf\_r  
DR Pfam; PF03552; Cellulose\_sy  
DR SMART; SM00184; RING; 1;  
DR PROSITE; PS01121; CASPASE  
SQ SEQUENCE 1065 AA; 120618  
Query Match 65.4%  
Best Local Similarity 64.1%  
Matches 701; Conservative 1  
QY 2 DGADALKSGRHGAGDVQCICA  
Db 16 DSDGDLKPLKDLNGCQCICG  
QY 62 LQCKTKYKRGSGPAIRGEGC  
Db 76 PCKARFRNGSPRVEDEKE  
QY 122 GNV-----GHPKYDSEIGE  
Db 133 ESLPSVLLTHGP-----  
QY 174 TGNISRAPPPY-----VNHS  
Db 162 C--IDPLPGIYQLLLPVRI  
QY 228 SIAPSEGAATDIDASTENNE  
Db 218 --KYHEKGGE--FEGTSGNG  
QY 288 IFLHYRLNPNVRNAYPLMLLSV  
Db 274 VELHYRTHPVKDAYALWLTSA  
QY 348 EPSQLAAVDIFVSTVDPLKEPE  
Db 334 EPSQLAPVDVSTVDPMKEPP  
QY 408 ETSEFARKWVPVKYDYTEPRA  
Db 394 ETAEFSKKWVPFCKKFNIEPR  
QY 468 NALVSKALKVPEEGWIMQDGP  
Db 454 NILVAKAKIPEDEGWTMEDGT  
QY 528 EKRPFGHKKAGAMNALVRV  
Db 514 EKRPFGHKKAGAMNALVRV  
QY 588 QVCYVQFPQRFQDIDRNDRYAN  
Db 574 KCCYVQFPQRFQDIDRNDRYAN

97; Shea T.P., Benito M.-I., Town C.D.,  
S.L., Barnstead M.E., Feldblum T.V.,  
J., Ronning C.M., Koo H., Moffat K.S.,  
E., Umayam L., Tallon L.J., Gill J.E.,  
T.H., Goodman H.M., Somerville C.R.,  
man W.C., White O., Eisen J.A.,  
er J.C.;  
osome 2 of the plant Arabidopsis  
/GenBank/DBJ databases.  
ynt.  
ore 3778.5; DB 10; Length 1065;  
ed. No. 3e-297;  
Mismatches 172; Indels 69; Gaps 14;  
TLGDGVFTACDVCRFPVRCYHEREGTQAC 61  
LTKTGNFVACNECGFPLCQCYEYERDGSQC 75  
DDGSDFNYPASSTEDOKOKIADRMRSWRMTGGS 121  
DIENEFDTQG---NNKARLPHRAEEFSSSSHE 132  
SGEIPRGVVPVSTNSQMSGEIPGASPDH--MMSP 173  
-----VSGEIP--TPDRNATLSP 161  
SF--SGSNGNVAWKEVDCWKKQDKGATPMTNGT 227  
DLNSYGLVNDWKKRIQCKMLKQDKNMLHMTG-- 217  
NDETRQPLSRKVPPIASSKINPYRMVIVLRLVLS 287  
DDARLPMSRVVHFPSPARMTYRIVIVLRLIILG 273  
AFALSWILDQFPKWPPIINRETYDLRLALRYDREG 347  
AFALSWILDQFPKWPPIINRETYDLRLALRYDREG 333  
NTVLSILAVDYPVDKVCVSDGASMLTFDALA 407  
NTVLSILAVDYPVDKVCVSDGASMLTFDALA 393  
FCQKIDYLDKQVPQFVDRRRAMKREYEEFKIRI 467  
FCQKIDYLDKQVPQFVDRRRAMKREYEEFKIRI 453  
FTRDPHGMQVFLHSGGLDTSGNELPRLVYVSR 527  
FTRDPHGMQVFLHSGGLDTSGNELPRLVYVSR 513  
NGOYMLNLDCHYINNSKAVREAMCFMDPNLGP 587  
NGOYMLNLDCHYINNSKAVREAMCFMDPNLGP 573  
FFDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPP 647  
FFDINLRGLDGIQGPVYVGTGCVFNRTAIYGEPP 633

```
QY 648 PIKAK--KPGFLASLCGGKKKASKK-RSSDKKSKNKHVDSSVPFNLEDIEGVEGAG 704
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 634 VLTEEDLENIIVKSCFGSKKSKRKIPNYEDNRSIKRSDSNVPLENMDIDEDVE--G 691
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 705 FDEKSVLMSQMSLEKRFQOSAAFVASTLMEYGGVQSQSTPESLLKEAIIHVISCYEDKS 764
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 692 YEDMSLLVSQRLEKRFQSQPVFAATFMEOGGLPSTNPLTLTLLKEAIIHVISCYEAKT 751
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 765 EMGTEIGWIYGSVTEIILTGFKNHARGWRSVYCMKRPAPKGSAPINLSDRLNQVLRWAL 824
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 752 DMGKEIGWIYGSVTEIILTGFKNHARGWISYCVPSRPAFKGSAPINLSDRLNQVLRWAL 811
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 825 GSVEILFSRCHPLWYGYGGRKLFERFAYINTIYPLTSLPLLYCYILPAICLTGKFIM 884
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 812 GSTEILLSRHCPITWYGYNGRLKLERIAYINTIVYPTSIPLLAYCMLPAFCLITNTFII 871
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 885 PEISNLASTWFIATFLSIFATGILEMRWSGVGIDEMWRNEOFWVIGGISAHLEAVFOGLL 944
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 872 PEISNLASLCFLLFASIYASALLEKWSDALEKWDNRNEQFWVIGGSAHLFAVFOGLL 931
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 945 KVLAGIDTNTVTSKANDEGDFAEIYFKWTTLLIPPTTILINMVGWVAGTSYAINSG 1004
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 932 KVFAGIDTNTVTSKASDEGDFAEIYVFKWTSLLIPTTILLYNLVGIVAGVSYAINSG 991
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1005 YQSWGPLFGKLFPAFWVIVHLYPFLKGLMGQRNRPPTIVVWAVLLASIFSLWVRVDPF 1064
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 992 YQSWGPLMGKLLFAFWVAHLYPFLKGLGRNRPPTIVVWSALLASIFSLWVRINPF 1051
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1065 TTRLAGPNIOTCGI 1078
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1052 -----VSTTGV 1057
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: February 19, 2003, 16:36:55  
Job time : 63 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2003, 16:34:12 ; Search time 25 Seconds  
(without alignments)  
1271.070 Million cell updates/sec

Title: US-09-900-237-30

Perfect score: 5778

Sequence: 1 MDGDADALKSRHAGDVCQ.....VDPFTRLAGPNIQTCGINC 1080

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4455.5	77.1	881	4	US-08-960-048-8
2	3477	60.2	974	4	US-08-960-048-6
3	2811	48.7	685	4	US-08-960-048-7
4	268	4.6	693	4	US-08-960-048-11
5	255.5	4.4	861	4	US-08-960-048-12
6	238.5	4.1	3031	1	US-07-689-008-2
7	237.5	4.1	756	4	US-08-960-048-10
8	237.5	4.1	756	4	US-08-960-048-3
9	207	3.6	723	4	US-08-960-048-9
10	207	3.6	723	5	PCT-US91-01726-4
11	114	2.0	357	1	US-08-119-773-4
12	112	1.9	1957	4	US-08-669-656A-8
13	111	1.9	1957	4	US-08-669-656A-2
14	108	1.9	1167	1	US-08-485-568A-6
15	108	1.9	1167	2	US-08-590-554A-6
16	108	1.9	1167	2	US-08-184-223-6
17	108	1.9	1168	1	US-08-620-717A-9
18	107.5	1.9	2132	4	US-08-669-656A-6
19	106	1.8	357	1	US-08-119-773-6
20	105	1.8	357	1	US-08-119-773-5
21	105	1.8	357	1	US-08-119-773-2
22	104.5	1.8	3033	1	US-07-925-695-5
23	102	1.8	616	4	US-09-136-574A-47
24	101	1.7	418	4	US-09-134-001C-4051
25	101	1.7	1026	1	US-07-998-003A-95
26	101	1.7	1026	1	US-08-453-274B-95
27	101	1.7	1026	1	US-08-453-695A-95

28	101	1.7	1026	1	US-08-268-161A-95	Sequence 95, Appl
29	101	1.7	1026	1	US-08-453-702A-95	Sequence 95, Appl
30	101	1.7	1026	4	US-09-099-639-95	Sequence 95, Appl
31	101	1.7	1026	5	PCT-US93-12588-95	Sequence 95, Appl
32	101	1.7	1026	5	PCT-US95-08071-95	Sequence 95, Appl
33	101	1.7	1203	1	US-07-998-003A-103	Sequence 103, App
34	101	1.7	1203	1	US-08-453-274B-103	Sequence 103, App
35	101	1.7	1203	1	US-08-453-695A-103	Sequence 103, App
36	101	1.7	1203	1	US-08-268-161A-103	Sequence 103, App
37	101	1.7	1203	2	US-08-453-702A-103	Sequence 103, App
38	101	1.7	1203	4	US-09-099-639-103	Sequence 103, App
39	101	1.7	1203	5	PCT-US93-12588-103	Sequence 103, App
40	101	1.7	1203	5	PCT-US95-08071-103	Sequence 103, App
41	100.5	1.7	1956	4	US-08-843-417-2	Sequence 2, Appl
42	99.5	1.7	2206	1	US-07-852-260-2	Sequence 2, Appl
43	99.5	1.7	2206	2	US-08-461-503-2	Sequence 2, Appl
44	99.5	1.7	2206	4	US-08-465-250-2	Sequence 2, Appl
45	98.5	1.7	3033	1	US-07-925-695-9	Sequence 9, Appl

#### ALIGNMENTS

RESULT 1

US-08-960-048-8  
; Sequence 8, Application US/08960048C

; Patent No. 6271443

; GENERAL INFORMATION:

; APPLICANT: Stalker, D. et al.

; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter

; FILE OF INVENTION: Sequences

; CURRENT APPLICATION NUMBER: US/08/960,048C

; CURRENT FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987

; PRIOR FILING DATE: 1996-10-29

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 8

; LENGTH: 881

; TYPE: PRT

; ORGANISM: Oryzae sativa

US-08-960-048-8

Query Match 77.1%; Score 4455.5; DB 4; Length 881;  
Best Local Similarity 93.4%; Pred No. 0;  
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;

QY	201	GNVAKKRVGKWKQDKGAIPMTNGTSIAPSEGRAATIDASTEYNMEDALLNDETROP	260
DB	1	GNVAKKRVGKWKQDKGAIPMTNGTSIAPSEGRGVGDIDASTDYNNEDALLNDETROP	60
QY	261	LSRKVPFIASSKINPYRMVIVLVLVLSIFLHYRLTNPNRAYPLWLISVCEIWFALSWI	320
DB	61	LSRKVPFIASSKINPYRMVIVLVLVLSIFLHYRLTNPNRAYPLWLISVCEIWFALSWI	120
QY	321	LDQFPKFPINRETYLDRLALRYDREGPEPSOLAAYDIFVSTVDPLKEPPIVTANTVLSIL	380
DB	121	LDQFPKFPINRETYLDRLALRYDREGPEPSOLAAYDIFVSTVDPLKEPPIVTANTVLSIL	180
QY	381	AVDYVDPVKVSCVSDDGASMLTFDALAETSEFARKWPFVKYKIDIEPRAPFVCKIDY	440
DB	181	AVDYVDPVKVSCVSDDGASMLTFDALAETSEFARKWPFVKYKIDIEPRAPFVCKIDY	240
QY	441	LKDKVQPSFVKDRRAMKREYEEFKIRINALVSKALYPEEGWIMQDGTWPGNNTRDHPG	500
DB	241	LKDKVQPSFVKDRRAMKREYEEFKIRINALVSKALYPEEGWIMQDGTWPGNNTRDHPG	300
QY	501	MIQVFLHSGGLDTEGNEPLRLVYSREKRPFGFQHHKAGAMNALVRVSAVLNNGOYMLN	560
DB	301	MIQVFLHSGGLDTEGNEPLRLVYSREKRPFGFQHHKAGAMNALVRVSAVLNNGOYMLN	360
QY	561	LDGDHYINNSKAVREAMCFLMDPNLGPQVCVQPPQRFQDGLDRNDRYANRNTVFFDINLR	620

Db 361 LDCDHYNNSKALREACFLMDPNLGRSVYQVFPQRFQDIDRNDRYANRNTVFEDINUR 420  
QY 621 GLDGIQGVYVGGVFNRTAIYGYEPPKAKKPG-FLASLCGGKKKASKKRSDDKK 679  
Db 421 GLDGIQGVYVGGVFNRTAIYGYEPPKAKKPG-FLASLCGGKKKASKKRSDDKK 480  
QY 680 SNKHVDSVPVFNLEDIEEGVEGAGFDDDEKSVLMSQMSLEKRFQSGAAAFVASTLMEYGV 739  
Db 481 SNKHVDSVPVFNLEDIEEGVEGAGFDDDEKSVLMSQMSLEKRFQSGAAAFVASTLMEYGV 540  
QY 740 POSTPESILLKRAHIVSCGYEDKSEMGEIGWISVTEIDLTGFKMHARGWRSVYCHP 799  
Db 541 POSTPESILLKRAHIVSCGYEDKSEMGEIGWISVTEIDLTGFKMHARGWRSVYCHP 600  
QY 800 KRPAFGSAPINLSDRLNOVLWALGSVEILFSRHCPWYGGRLKFLERFAYINTTLY 859  
Db 601 KRPAFGSAPINLSDRLNOVLWALGSVEILFSRHCPWYGGRLKFLERFAYINTTLY 660  
QY 860 PLTSLPLLYCILPAICLTGTFIMPEISNLSINFIALFSLFATGILEMRWSGVGIDE 919  
Db 661 PLTSLPLLYCILPAICLTGTFIMPEISNLSINFIALFSLFATGILEMRWSGVGIDE 720  
QY 920 WRNEQFWYIGTSAHLFAVFGGLKVLZAGIDTFTVTYSKANDDEGDFAEIYFKWTLL 979  
Db 721 WRNEQFWYIGTSAHLFAVFGGLKVLZAGIDTFTVTYSKANDDEGDFAEIYFKWTLL 780  
QY 980 IPTTILINMGVAVAGTSYAINSGYQSGWGLFGLKFFAFVIVHLYPPLKGLMGQRNKT 1039  
Db 781 IPTTILINMGVAVAGTSYAINSGYQSGWGLFGLKFFAFVIVHLYPPLKGLMGQRNKT 840  
QY 1040 PTIVVAVILLASIFSLLMVRVDPFTTRLAGPNIQTCGNC 1080  
Db 841 PTIVVAVILLASIFSLLMVRVDPFTTRLAGPNIQTCGNC 881

RESULT 2  
US-08-960-048-6  
; Sequence 6, Application US/08960048C  
; Patent No. 6271443  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: Sequences  
; CURRENT APPLICATION NUMBER: US/08/960,048C  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR FILING DATE: 1996-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 974  
; TYPE: PRT  
; ORGANISM: Gossypium hirsutum  
US-08-960-048-6

Query Match 60.2%; Score 3477; DB 4; Length 974;  
Best Local Similarity 61.5%; Pred No. 1.9e-318;  
Matches 657; Conservative 129; Mismatches 176; Indels 106; Gaps 15;

QY 18 VCOICADGLGTLDDGVFTACDVCRRPVCYHERKEGTQACQCKTKYKRHRGSPA 77  
Db 8 VCHTCGERHVLNVNPPFVACHCENFPICKSCFEYDLKGRKACLCR-----GSPY- 58  
QY 78 RGEEDGDDADGSDNFYASCTEDQKQKIADMRMRWMTGGSGNVG-HPKYDSGEIGL 136  
Db 59 -----DENLDD-----VEKATGQOSTMAA-----HLNKSQDVGIHARHIS---SV 96  
QY 137 SKYDSGEIPRGVPSVTSNQSCEIPGASPDHMHMSPTGNISRRAPFPVNHSPNPSREF 196  
Db 97 STLDS-----EM 103

QY 197 SGSGINVAWKERVDCGKMKQDK 420  
Db 104 AEDNGNSIWKRVESWKEKKN 154  
QY 257 TRQPLSRKVPYASSKINPYRMV 316  
Db 155 S-OPLSTIIPKSLAPYRTV 213  
QY 317 LSWILDQPKWFPNRETYLDR 376  
Db 214 FSWILDQPKWFPNRETYLDR 273  
QY 377 LSILAVDPYDKVSCYVSDDA 436  
Db 274 LSILALDPYDKVSCYVSDDA 333  
QY 437 KIDYLDKQVQSFVKRRAMKR 496  
Db 334 KIDYLDKQVQSFVKRRAMKR 393  
QY 497 DHPGMIQVFLGSHGGLDTGNE 556  
Db 394 DHPGMIQVFLGSHGGLDTGNE 453  
QY 557 YMLNLDCHYINNSKAVREAM 616  
Db 454 FTLNLDCHYINNSKAVREAM 513  
QY 617 INLRGLDGIQGVYVGTGVFN 676  
Db 514 VNMGLDGIQGVYVGTGVFN 573  
QY 677 KKSNNKHVDSVPVFNLEDIE 735  
Db 574 LYRDAKREELDAIFNLREID 627  
QY 736 YGGVQPSPTESILLKEAIVH 795  
Db 628 NGVAESANPSTLKEAIVH 687  
QY 796 YCMPKRPAPKGSAPINLSDRL 854  
Db 688 YCMPKRPAPKGSAPINLSDRL 747  
QY 855 NTTPPLTSLPLLYCILPAIC 914  
Db 748 NTTPPLTSLPLLYCILPAIC 807  
QY 915 VGIDWWRNEQFWIGGISAH 974  
Db 808 VSIEDLWRNEQFWIGGISAH 866  
QY 975 WTTLLIPPTTLLIINMGVVA 1034  
Db 867 WTTLLIPPTTLLIINMGVVA 926  
QY 1035 QONRTPTIIVVAVILLASIF 1080  
Db 927 QONRTPTIIVVAVILLASIF 974

RESULT 3  
US-08-960-048-7  
; Sequence 7, Application US/08960048C  
; Patent No. 6271443  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: Sequences  
; CURRENT APPLICATION NUMBER: US/08/960,048C  
; PRIOR FILING DATE: 1997-10-29  
; PRIOR APPLICATION NUMBER: 6271443

; PRIOR FILLING DATE: 1996-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Gossypium hirsutum  
US-08-960-048-7

Query Match 48.7%; Score 2811; DB 4; Length 685;  
Best Local Similarity 74.1%; Pred. No. 6.1e-256;  
Matches 509; Conservative 78; Mismatches 76; Indels 24; Gaps 6;  
QY 413 ARKWPFFVKKYDIETPRAPFEYFCOKIDYLDKQVQSFYKDRAMKREYEEFKIRINALVS 472  
Db 1 ARRWPFCKKHNVPEPAPEYFNEKIDYLDKQVHPFSYKERRAMKREYEEFKIRINALVA 60  
QY 473 KALKVPEEGWIMODGTPWGNTRDPGMIOVFLGHSGLDTEGNEPLRLVYVSREKRP 532  
Db 61 KAQKKPEEGWIMODGTPWGNTRDPGMIOVFLGHSGLDTEGNEPLRLVYVSREKRP 120  
QY 533 FOHKKAGAMNALVRVSALVTNGQYMLNDCDHYINNSKAVREAMCFMDPDLGQVQCYV 592  
Db 121 YOHHKAGAMNALVRVSALVTNAPFTLNDCDHYINNSKAVREAMCFMDPDLGQVQCYV 180  
QY 593 QFQRTDGDIDRNDRYANRNVTFEDINLRGLDGIQGVYVGTGCVENRTAIYGYEPIKAK 652  
Db 181 QFQRTDGDIDRNDRYANRNVTFEDINLRGLDGIQGVYVGTGCVENRTAIYGYEPIKAK 240  
QY 653 K-----PGFLASLCGKKKSKKRS-----SDKKK---SNKHVDSSVPVFN 692  
Db 241 RPKMTDCWPSCWCCCGGSRKSKKKKGLLGLLGGKMMKMMKMYKKGSAPVFD 300  
QY 693 LEDIEEGVEGAGDD--EKSVLMSQMSLEKRFQGSAAFAVASTIMEYGGVQPSSTPESLKE 751  
Db 301 LEEIEEGLE--GYEELEKSTLMSQKNEKRFQGSAPVFIASTLMENGGLPEGTNSTSLKE 358  
QY 752 AIHVISCGYBDKSEWTEIGWIGSVTEIDLTGFKMHARGWSVCMKRPAPKFSAPIN 811  
Db 359 AIHVISCGYEKTEWGEIGWIGSVTEIDLTGFKMHARGWSVCMKRPAPKFSAPIN 418  
QY 812 LSRLNOVLKALGSEVILFRRHCPHLYGGRKLEFARFAYINTIYPLTSLPLLYYCI 871  
Db 419 LSRLHQLKALGSEVILFRRHCPHLYGGRKLEFARFAYINTIYPLTSLPLLYYCI 478  
QY 872 LPAICLLTGKFIIMPEISNLASIFIALFLSIFATGILEMRWSGVGTDEWRNEQFWVIGG 931  
Db 479 IPAVCLLTGKFIIPTLSNLTSSVWFLALFLSIFATGILEMRWSGVSIQDWRNEQFWVIGG 538  
QY 932 ISAHFAVFOGLKLVLAGIDTNTFTVTSKANDEEGDFAELMFKWTLTLLIPTTILINMV 991  
Db 539 VSAHFAVFOGLKLVLAGIDTNTFTVTSKANDEEGDFAELMFKWTLTLLIPTTILINMV 597  
QY 992 GVVAGSYANSGYSGWGLFGLKFFAFWVILHLPFLKGLMGRQNRPTIIVVAVLLA 1051  
Db 598 GVVAGVSDALNNGYSGWGLFGLKFFAFWVILHLPFLKGLMGRQNRPTIIVVAVLLA 657  
QY 1052 SIFSLWVRVDPFTTTLRAGPNIQTCGI 1078  
Db 658 SIFSLWVRIDPFLPKQTPVLKQCGV 684

RESULT 4  
US-08-960-048-11  
; Sequence 11, Application US/08960048C  
; Patent No. 6271443  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: 15621/01/US  
; CURRENT APPLICATION NUMBER: US/08/960,048C  
; CURRENT FILLING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987  
; PRIOR FILLING DATE: 1996-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11  
; LENGTH: 693  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-08-960-048-11

Query Match 4.6%; Score 268; DB 4; Length 693;  
Best Local Similarity 19.5%; Pred. No. 4.1e-16;  
Matches 156; Conservative 83; Mismatches 201; Indels 360; Gaps 29;  
QY 263 RKVPFIASINPYRMVIVLRVLVLSIFLHYRLT-----NPVRNAYPLMLLSVICSIFA 316  
Db 11 RRMGPRESAL-----MLIVLSLVTSVCRIWMRYTSTLNWDDPVSGLGILLFAITAVIV 66  
QY 317 LSWILDDQFPKWFPIINRETYLDRLALRYDRGEPESQLAAVDIFVSTYDPLKEPPIVTANTV 376  
Db 67 L--VLGYFQVYVPLNRQP-----VPLPKDMSLWPS-----VDIFVPTYN---EDUNVVKNTI 113  
QY 377 LSLIADVPYDKVSCYVSDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPEYFCQ 436  
Db 114 YASIGIDWPDKLNIWLDGG----- 135  
QY 437 KIDYLDKQVPSFKVDRRAMKREYEEFKIRINALVSALKLVPEEGWIMODGTPWPGNNTR 496  
Db 136 -----REEFRQFQNVG----- 147  
QY 497 DHPGMIQVFLHSGGLDTEGNEPLRLVYVSREKRPFGQHHKAGAMNALVRVSALVTNGQ 556  
Db 148 -----VKYIARTT-----HEHAKAGNINNAKYA---KGE 174  
QY 557 YMLNDCDHYINNSKAVREAMCFMDPDLGQVQCYVQFQRF---DGIDRN-----DRYAN 609  
Db 175 FVSTFDCDHYVTPRSFLQMTGWFELKE-----KOLAMQTPHHFSPDPFERNLGRFRKTPN 230  
QY 610 RNTVFEDINLRGLDGIQGVYVGTGCVENRTAIYGYEPIKAKKPGFLASLCGKKKASK 669  
Db 231 EGTUFLYGLVODGNDMDATFFCGSCAVIR-----Kp----- 262  
QY 670 SKRSSDKKSNKHVDSSVPVFNLEIDTEGVEGAGFDDKSVLMSQMSLEKRFQGSAAFV 729  
Db 263 ----- 262  
QY 730 ASTLMYEGGVQPSSTPESLKEAHLVISCYEDKSEWTEIGWIGSVTEIDLTGFKMHA 789  
Db 263 ---LDEIGGI-----AVE-----TVTDAHTSURLHR 286  
QY 790 RGMRSVCMKRPAPKFSAPINLSDRNLQVLRWALGSVEILFRRHCPHLYGGRKLEF 849  
Db 287 RGYTSAYM--RIFOAAGLATESLAHIGORIRWARGVQI--FRLDNPL---TGKGLKFAQ 340  
QY 850 RFAYINTIYPLTSLPLLYYCIIP-----AICLLTGKFIIP----- 885  
Db 341 RLCYVNAFMHFLSGIPRLIFLTAPLAFLLHAYIITAPALMIALFVLPHMIHASLTNSKI 400  
QY 886 -----EISNLASIFIA-----LFLSIFATGILEMRWSGVGTDEWRNEQF--WVI 929  
Db 401 QGKYRHSFWEIYETVLAWYIAPTTLVALLNPHKGFENVTAKGGGLVE-----EYVDWVI 456  
QY 930 GGISAHLFAVFOGLKLVLAGIDTNTFTVTSKANDEEGDFAELMFKWTLTLLIPTTIL--I 987  
Db 457 S--RPYIFLVLNLLVGVAVGI-----WRYFVGPPTMLTVV 490  
QY 988 INWVGV-----VAGTSYAIN 1002  
Db 491 VSMVWVFNILVLGGAVAVS 510

RESULT 5  
US-08-960-048-12

NUCLEIC ACID SEQUENCES FOR THE  
OF CELLULOSE SYNTHASE OPERON

```
Db 191 -----VRP-----EFGQFADCGAL----- 205
QY 496 RDHPGMIQVFLGHSGGLDTGEGNELPRLVVYSREKRGFGQHHKKAGAMNALVRVSAVLING 555
Db 206 -----YIGRVDS-----HAKAGNLNHAIK-----RTSG 229
QY 556 QYMLNDCDHYINNSKAVREAMCFMLDPNLGPQVCYVQFPQFDGIDRNDRYANRNTVFF 615
Db 230 DYILILDCDH-IPTRAFLOIAMGMWVADR---KIALMQTPHHFYSPD----- 272
QY 616 DINLRGLDGTQGPVYVGTGCFVNRNTAIYGEYEPPIKAKKPGFLASLCGGKKKASKKSS 675
Db 273 -----PFORNLAVGRTD----- 285
QY 676 DKKSNKHVDSSVPFVNLEIDIEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFAVASTLME 735
Db 286 -----PEGNL---FYGVIQDGN-----FWDATFFCGSCAI- 313
QY 736 YGVPQSPSPESLLKEAIIHVISGCEYEDKSEWTEIGWYIGSVTEIDLTGFKMHARGWRSV 795
Db 314 -----LREAAIESIG-----GFVETVTEDAHTALRMQRGWSTA 348
QY 796 YCMPKRPAPKGSAPINLSDRNLQVLRWALGSEVILSRHCPWYGGRLKFLERFAYIN 855
Db 349 YL--RIPVASGLATERLTHIGORMWARGMIOI-FRVDNPM---LGGGLKGLQRLCYLS 402
QY 856 TTYPLTSLPLLYCYLPAICLTGKFIPEISNLASIFIALFSIFATGILEMRS-- 913
Db 403 AMTSEFFAIPRVIFLASPLAFPFQGNII-----AASPLAVLAYAIPHPHSTA 451
QY 914 -GVGIDSWNEQFW-VIGGISAHLEFAVFGLLKVLGIDTNTFTVTSKA---NDEGDPA 968
Db 452 TAAKYNKGNR-YSEWSEVETTWALFLVRVTIITLFPKSGKFNTEKGVLEEFDFLG 510
QY 969 ELYMFKWTTLLIPTTILIIINMVGWVAGTSYAINSGYQSGWPLFG--KLFFAFWVIVHLY 1026
Db 511 ATY-----PNIIF-----AGIMTLGLLIGLFEITFHF----- 537
QY 1027 PFLKGLMGRONRPTIVVWA-----VILLASI 1053
Db 538 ---NQLAGIAKRAYLNCIWMASLILLAAI 566

RESULT 7
US-08-960-048-10
; Sequence 10, Application US/08960048C
; Patent No. 6271443
; GENERAL INFORMATION:
; APPLICANT: Stalker, D. et al.
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter
; FILE REFERENCE: 15621/01/US
; CURRENT APPLICATION NUMBER: US/08/960,048C
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/029,987
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 756
; TYPE: PRT
; ORGANISM: acetobacter xylinum
US-08-960-048-10

Query Match 4.1%; Score 237.5; DB 4; Length 756;
Best Local Similarity 20.1%; Pred. No. 3 5e-13;
Matches 156; Conservative 89; Mismatches 232; Indels 301; Gaps 31;

QY 280 VLRVLSIFLHYRLNPNVR-NAYPLWLLSVI---CEINWALSIIWLDQFPKWFPPINRETY 335
Db 80 VLSALVSLRYLTWRLTETLDNFNWIQIGLIVLLMAELVLYLFLSYFOTIQLHRAP- 138
```

```
QY 336 LDRLALRYDRGEPSQAAVDIFVSTVDPLKEPPIVITANTVLSILAVDPVSKVCSYSD 395
Db 139 ---LPLPDNDWDWT---VDIPTVD---EQLSIVRLVGLGIDWPPDKVNVYILD 188
QY 396 DGASMLTFDALAETSEFARKWVFVKYDIEPRAPEFYFCOKIDYLDKXQVPSFVKDRA 455
Db 189 DG-----VRP----- 193
QY 456 MKREYEFKIRINALVSKALVPEEGWIMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTE 515
Db 194 ---EFGQFADCGAL-----YIGRVDD 213
QY 516 GNEPLRVLYVSREKRGFGQHHKKAGAMNALVRVSAVLNGOYMLNDCDHYINNSKAVRE 575
Db 214 -----SAHAKAGNLNHAIK-----RTSGDYILILDCDH-IPTRAFLOI 250
QY 576 AMCFMLDPNLGPQVCYVQFPQFDGIDRNDRYANRNTVFFDINLRGLDGTQGPVYVGTGC 635
Db 251 AMGMWADR---KIALMQTPHHFYSPD----- 274
QY 636 VFNRTAIYGEYEPPIKAKKPGFLASLCGGKKKASKKRSDDKKSNKHVDSSVPFVNLE 695
Db 275 PFORNLAVGRTD-----PEGNL-- 292
QY 696 IEEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFAVASTLMEYGGVPOSSPESLLKEAIIH 755
Db 293 -FYGVIQDGN-----FWDATFFCGSCAI-----LREAAIES 323
QY 756 ISCYEDKSEWTEIGWYIGSVTEIDLTGFKMHARGWRSVYCMKRPAPKGSAPINLSDR 815
Db 324 IG-----GFVETVTEDAHTALRMQRGWSTAYL--RIPVASGLATERLTH 368
QY 816 LNOVLRWALGSEVILSRHCPWYGGRLKFLERFAYINTTYPLTSLPLLYCYLPAI 875
Db 369 IGQRMWARGMIOI-FRVDNPM---LGRGLKGLQRLCYLSAMTSEFFAIPRVIFLASPLA 434
QY 876 CLLGKGTIMPRISNLASIFIALFSIFATGILEMRS---GVGIDSWNEQFW-VIGG 931
Db 425 FLFAGQNTII-----AAPLAVAYALPHMFHSIATAAKVNKGNR-YSEWSEVYE 472
QY 932 ISAHLEFAVFGLLKVLGIDTNTFTVTSKA---NDEGDFAELY-MFKWTTLLIPTPTI-- 985
Db 473 TTMALFLVRVTITVTLFPKSGKFNTEKGVLEEFDFLGATYPTNIIFATIMMGGLLIGL 532
QY 986 --LII--NMVGWVAGTSYAINSGYQSGWPLFGKLFFAFWVIVHLYPELKL-MGRONR 1038
Db 533 FELIVRFNQLDVIARNAYLLNCA-----WALISLIIILFAAIVGRETK 575

RESULT 8
US-09-147-236-3
; Sequence 3, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiro
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-0PCT
; CURRENT APPLICATION NUMBER: US/09/147,236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 756
; TYPE: PRT
; ORGANISM: Acetobacter xylinum
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
```







OTHER INFORMATION: /note= "The G-6-Pase amino acid sequence mutated from Arg to Cys at position 83."  
US-08-119-773-4

Query Match 2.0%; Score 114; DB 1; Length 357;  
Best Local Similarity 19.1%; Pred. No. 0.043;  
Matches 72; Conservative 59; Mismatches 137; Indels 108; Gaps 16;  
Qy 751 EAHVISGDEKSWGTEIGWISGVTEDILTGFKMHARGWRSVYCMKPRPAFKGSAPI 810  
Db 14 QSTHYLVNQYQSDW-----FILSVIADLRNAPYVLPFW---FHLQEAVGIKLLWA 65  
Qy 811 NLSDRNLQVLRWALGSVILSRHCPKLVWGGGRKLFERAYINTTPTTSLDLL--- 867  
Db 66 VIGDWLNLVFKW-----ILFQ-CPYWM-----VLDTDYISNTSVPLIKQF 105  
Qy 868 -----VYCIU-PAICLLTGKFIIMPEIS----- 888  
Db 106 PVTCTGPGSPSGHAMGTAGYVYVNVVTSLSIFQCK-IPKYRFRCLNVILWLGFWAYOL 164  
Qy 889 --NLASIWFIALFISFATGILEMRWSGVGIDEMWRN-----EOWVWIGGISAHLF 937  
Db 165 NVCLSRIVIAAHFPHQVAVGL---SGIAVTETESHISYNASLKKYFL---ITFFLF 217  
Qy 938 AVFQGLLVLAGIDNFTVTSKANDDEGDEAFELYMFKWTTLLIPPTTILIIIMVGVVAGT 997  
Db 218 SFAIGFYLLKGLGVLDLTLWLEKAGRWCEQPE-----VWHDITPPFASLLKLNL-GTLFGL 271  
Qy 998 STAINSG-----YOSWGPFLFKLFFAFWVHLVFPFLKGLMGRNRTTIVVWAV 1048  
Db 272 GIALNSMYRESCKGLSKWLPFRLLSSIVASLVLLHVEDSLK---PPSQVELVYVLSF 327  
Qy 1049 LLASIFLLWVRVDPF 1064  
Db 328 CKSAVVPLASVSVIPI 343

RESULT 12

US-08-669-656A-8  
Sequence 8, Application US/08669656A  
Patent No. 6451554

GENERAL INFORMATION:

APPLICANT: Wood, John N.  
APPLICANT: Akopian, Armen N.  
TITLE OF INVENTION: Ion Channel  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESS: ZENECA Pharmaceuticals  
STREET: 1800 Concord Pike, P.O. Box 15437  
CITY: Wilmington  
STATE: Delaware  
COUNTRY: USA  
ZIP: 19850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/669,656A

FILING DATE: 24-JUN-1996

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Hohenschutz, Liza D.

REGISTRATION NUMBER: 33,712

REFERENCE/DOCKET NUMBER: PHM.70086

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 886-7466

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 1957 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein  
us-08-669-656A-8  
Query Match 1.9%;  
Best Local Similarity 17.5%;  
Matches 207; Conservative 177; Indels 108; Gaps 16;  
Qy 66 TKYRHRGSPAIRGEEDG--- 121  
Db 439 TSLQSHSGSPASKANERRE 490  
Qy 121 -----SCNVCHPKYDSC 162  
Db 491 LSSGRRRASHGVSFHFRAPOS 550  
Qy 163 GASPDHMHMSPTGNISRRAP 220  
Db 551 -----RSPLE 571  
Qy 221 IPMTNGTSIAPSEGRATDIDA 279  
Db 572 VPTGELTAGAP-EGPAL--- 609  
Qy 280 VLRVWLISIFLHYRLTNPNR 338  
Db 610 AMSVSI-----MTSVIE 655  
Qy 339 LALRYDREGSPSQAADVIFV 395  
Db 656 MAL-----FE 691  
Qy 396 DGASMLTFDALAE----- 736  
Db 692 -----AFDAMLQAGNIVTV 779  
Qy 446 QPSFVKDRAMKREYEEFKIR 505  
Db 737 VTVSLLLELSASKG----- 779  
Qy 506 LGHSGGLDTEGNELPRLVYVS 560  
Db 780 IGNS----- 810  
Qy 561 LD----- 579  
Db 811 EDYGRKDGVSVMNGEKLWHE 870  
Qy 580 LMDPNLGPQVCYVOPQRFQ 628  
Db 871 LTVNVLGNLVNLNLF----- 912  
Qy 629 -----VYVGC 912  
Db 913 ALARIQVLGHRASRAIASYIS 972  
Qy 662 GKKKASKKRSDDKKKNK 709  
Db 973 NLTPALSSPKENHGDFTDP 1032  
Qy 710 -----SVL 741  
Db 1033 EQLPQVKCNHQAARSPAS 1092  
Qy 742 SSTPEILLKEAHHVI----- 781  
Db 1093 CPDPEELRKIPPELADLDEF 1144  
Qy 782 LTGFKMHARGWRSVYCMKPR 836  
Db 1145 KTCYRIVEHSWFESFIIFM 1204  
Qy 837 L-WYGYGGRUKFLERFAYINT 880





```
;
; FILING DATE: 21-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/918,345
; FILING DATE: 21-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/558,738
; FILING DATE: 27-JUL-1990
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,698
; FILING DATE: 16-DEC-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,403
; FILING DATE: 30-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/999,053
; FILING DATE: 14-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA48DD2.C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-590-554A-6

Query Match      1.9%; Score 108; DB 2; Length 1167;
Best Local Similarity 19.7%; Pred. No. 1.2;
Matches 139; Conservative 77; Mismatches 214; Indels 276; Gaps 36;

QY 411 EFARKW-----VPFK--KYDIEPRAPEFYFCQKIDYLDKQVPSFKVDRRAKREYEE 462
Db 250 QFAERKWSVYDESFINTQKVDLQRIQD--YSTVTSTTFEKFETLNPSSKESVKNYNR 307
QY 463 FKIRINALYSKALV-----PPEGWIMODGT-----PWPGNNTRDHPGM 501
Db 308 Y---VRSMTLQSLDIAATWPTLDNVPNSVDIQLDQRLVFSVDVAGPWEGNDNIT-SNI 363
QY 502 IQVFELHSGGLD-TEGNEPLRLVYVSREKRGFGQHHKAGAMNALRVSAVLNCOYMLN 560
Db 364 IDVLTPINTGIGFQESSDLRKFY-PRIELQSMQFH-----GQYV-- 402
QY 561 LDCDHYINNSKAVREAMCFLMDPLNGPQVCYVQFPQRFQDIDRNDRYANR----- 610
Db 403 -----NSKSVEH-----CY-----SDGLKLN--YANKTITAGVSNID 432
QY 611 -----NTVPFDINLRGLDGIQGPVYVG-----TGCYFNRTAIYG 644
Db 433 ESNQNNKHNGYPVINSPTIDVNS-----QNSQYLDLNSVMVNGQKVTCG----- 479
QY 645 YEPPIKAKPGFLASLCGGKKKSKKSSDKKKSNKHD-----SSVPVFNLED 695
Db 480 --SPSSNGNSNAALPNQKINVIYS-VQSNDKPE--KHADTYRKWGYMSSHIP-YDL-- 531
QY 696 IEQVEGAGFDDKE--SVLMSQMSLEKRFQSAAFV-----ASTLMYEGGVQSPSTP 745
Db 532 VPENVIGDIDPTKQPSLLKGFPAEKGYGDSIAIVSEPLGANAVKLTSYQVLOMEVTN 591
QY 746 ESKLKEAIHV--ISCGYEDKSEWGEIGWYGSVTED-----ILTFKMHARGWRSYIC 797
Db 592 QTTQKYRIRIRVATGDTAASIWFHILIGPSGNDLTNEGHNFSSVSRNKMFEVQGNNGKY- 650
```

```
QY 798 MPKPAFKGSAPINLSRDLNQVLRWALGSVEILFSRHCPWLWYGYGRLKFLERFAYINTT 857
Db 651 -----VNLILTDSIELPSGQQTILIONTSQDLFLDRIEFIS-- 687
QY 858 IYPLTSLPLLVYCILPAICLLTGKFMPE-----ISNLASIWFIALFLSIFATGILEMRWS 913
Db 688 -LPSTSTP-----TSTNFVEPESLEKIINQVN---QLFSSSSQTEL----- 724
QY 914 GVGIDEMWRNEQFWVIGGISAHL---FAVFOQLLKVLAGIDNFTVTSKA----- 960
Db 725 -----AHTVSDYKIDQVVLKYNALSDDVFGVEKKALKRLVNAQAK 764
QY 961 -----NDEEGD-----FAELYMFKWTTLLIPPTTI 985
Db 765 LSKARNVLVGGNFEKGEHWALSREATMVANHELFGDHLPPPPTL 810
```

Search completed: February 19, 2003, 16:37:43  
Job time : 47 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 19, 2003, 16:36:16 ; Search time 31 Seconds  
(without alignments)  
890.090 Million cell updates/sec

Title: US-09-900-237-30  
 Perfect score: 5778  
 Sequence: 1 MDGDADALKSGRHAGDVQ.....VDFETRLAGPNIOTCGINC 1080

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues  
Total number of hits satisfying chosen parameters: 140259

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pcp.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length				
1	5778	100.0	1080	10	US-09-900-237-30	Sequence 30, Appl	
2	4507	78.0	1065	10	US-09-900-237-33	Sequence 33, Appl	
3	4455.5	77.1	881	10	US-09-838-539-8	Sequence 8, Appl	
4	4021.5	69.6	1091	10	US-09-900-237-26	Sequence 26, Appl	
5	3844.5	66.5	1165	10	US-09-900-237-8	Sequence 8, Appl	
6	3820.5	66.1	1039	10	US-09-900-237-14	Sequence 14, Appl	
7	3818	66.1	1086	10	US-09-900-237-10	Sequence 10, Appl	
8	3818	66.1	1148	10	US-09-900-237-4	Sequence 4, Appl	
9	3477	60.2	974	10	US-09-838-539-6	Sequence 6, Appl	
10	3404.5	58.9	701	10	US-09-900-237-32	Sequence 32, Appl	
11	3234	56.0	793	10	US-09-900-237-18	Sequence 18, Appl	
12	3038.5	52.6	740	10	US-09-900-237-24	Sequence 24, Appl	
13	2816	48.7	685	10	US-09-900-237-31	Sequence 31, Appl	
14	2811	48.7	685	10	US-09-838-539-7	Sequence 7, Appl	
15	2689	46.5	506	10	US-09-900-237-20	Sequence 20, Appl	
16	2627.5	45.5	610	10	US-09-900-237-16	Sequence 16, Appl	
17	2044.5	35.4	431	10	US-09-900-237-28	Sequence 28, Appl	
18	1418.5	24.6	320	10	US-09-900-237-6	Sequence 6, Appl	
19	1379.5	23.9	304	10	US-09-900-237-2	Sequence 2, Appl	

## ALIGNMENTS

```

RESULT 1
US-09-900-237-30
, Sequence 30, Application US/09900237
, Patent No. US20020120124A1
, GENERAL INFORMATION:
, APPLICANT: Allen, Stephen
, TITLE OF INVENTION: Plant Celluloses
, FILE REFERENCE: BB1170 US CIP
, CURRENT APPLICATION NUMBER: US/09/9
, CURRENT FILING DATE: 2001-07-06
, PRIOR APPLICATION NUMBER: 60/092,84
, PRIOR FILING DATE: 1998-07-14
, PRIOR APPLICATION NUMBER: PCT/US99/
, PRIOR FILING DATE: 1999-07-13
, PRIOR APPLICATION NUMBER: 09/720383
, PRIOR FILING DATE: 2000-12-21
, NUMBER OF SEQ ID NOS: 33
, SOFTWARE: Microsoft Office 97
, SEQ ID NO 30
, LENGTH: 1080
, TYPE: PRT
, ORGANISM: Triticum aestivum
US-09-900-237-30

```

	Query Match	100.0%;	Score 5778;	DB 10;	Length 1080;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 1080;	Conservative	0;	Mismatches	0; Indels
					0; Gaps
Qy	1	MDGDADALKSGRHGAGDVCQICADGLGTTLDGDVFTACDVCGRFVPCRPCEHERKEGTQA	60		
Db	1	MDGDADALKSGRHGAGDVCQICADGLGTTLDGDVFTACDVCGRFVPCRPCEHERKEGTQA	60		
Qy	61	CLOCKTKYKRRHGSPTARGEGDDTDADGSDFNYPASGTEDQKQTADRMRSWRMTGG	120		
Db	61	CLOCKTKYKRRHGSPTARGEGDDTDADGSDFNYPASGTEDQKQTADRMRSWRMTGG	120		
Qy	121	SGNVGHPKYDSGETGLSKYDSGEIPRGYPVSVTNQWSMGEIPGASPOHHMSPGTNISRR	180		
Db	121	SGNVGHPKYDSGETGLSKYDSGEIPRGYPVSVTNQWSMGEIPGASPOHHMSPGTNISRR	180		
Qy	181	APFFYVNHSPNPSREFSGSIGNVAKKRVDCGWMKQKQKGAIPMPTNGTISIAPSEGRAADI	240		

Db 181 APPYYNHSNPBSRESGSNGVAMKVRDGVKMKDKGAIPMTNGTSTAPSEGRAATDI 240  
QY 241 DASTEXNEDALLNDETROPLSRKVPFIASSKINPYRMWIVLRVLVLSIFLHYRLTNPVRN 300  
Db 241 DASTEXNEDALLNDETROPLSRKVPFIASSKINPYRMWIVLRVLVLSIFLHYRLTNPVRN 300  
QY 301 AYPLWLLSVCEIWFALSILWIDQPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFS 360  
Db 301 AYPLWLLSVCEIWFALSILWIDQPKWFPINRETYLDRLALRYDREGEPSQLAAVDIFS 360  
QY 361 TVDPLKEPPIVANTVLSILAVDYPVDKVSICYSDDGASMLTFDALAETSEFARKWVPFV 420  
Db 361 TVDPLKEPPIVANTVLSILAVDYPVDKVSICYSDDGASMLTFDALAETSEFARKWVPFV 420  
QY 421 KKYDIEPRAPEFVFCOKIDYLDKDVOPSFVKDRRAKREYEEFKIRINALVSKALVPEE 480  
Db 421 KKYDIEPRAPEFVFCOKIDYLDKDVOPSFVKDRRAKREYEEFKIRINALVSKALVPEE 480  
QY 481 GWIMQDGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPRLVYVVSREKRPFGQHHKKAG 540  
Db 481 GWIMQDGTWPNGNTRDHPGMIQVFLGHSGLDTEGNEPRLVYVVSREKRPFGQHHKKAG 540  
QY 541 ANNALVRVSAVLNTGQYMLNLDCHYINNSKAVREACMFMDPNLPQVCYVQFPORFDG 600  
Db 541 ANNALVRVSAVLNTGQYMLNLDCHYINNSKAVREACMFMDPNLPQVCYVQFPORFDG 600  
QY 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYBPPIKAKKPGFLASL 660  
Db 601 IDNRDRIANRNTVFFDINLRGLDGIQGPVYVGTGCVFNRTAIYGYBPPIKAKKPGFLASL 660  
QY 661 CGGKKKASKSKRSDDKSNKHDVSSVPVFNLEDEEGVEGAGFDDKSKVLSMSQMSLEK 720  
Db 661 CGGKKKASKSKRSDDKSNKHDVSSVPVFNLEDEEGVEGAGFDDKSKVLSMSQMSLEK 720  
QY 721 RFGQSAFVASTMEYGGVQPSQSTPESLKEAHLVISCYEDKSEWGTGWTGIVGVTED 780  
Db 721 RFGQSAFVASTMEYGGVQPSQSTPESLKEAHLVISCYEDKSEWGTGWTGIVGVTED 780  
QY 781 ILTGFKMHARGWSVVCMPKRPFAKGSAPINLSDRLNOVLRWALGSVEILLFSRHCPLWYG 840  
Db 781 ILTGFKMHARGWSVVCMPKRPFAKGSAPINLSDRLNOVLRWALGSVEILLFSRHCPLWYG 840  
QY 841 YGRLKFLERFAYINTIYPLTSLPLLVYCIILPAICLLTGKFTMPISNLIASIFIALFL 900  
Db 841 YGRLKFLERFAYINTIYPLTSLPLLVYCIILPAICLLTGKFTMPISNLIASIFIALFL 900  
QY 901 SIFATGILEMRWSGVGIDENWNEQFWIGGISAHLPAVFGQLLKVLAGIDTFTVTSKA 960  
Db 901 SIFATGILEMRWSGVGIDENWNEQFWIGGISAHLPAVFGQLLKVLAGIDTFTVTSKA 960  
QY 961 NDEEGDFAELYMEKWTLLIPPTTILINMGVVAGTSVAINSGYOSWGPLFGKLPFAFW 1020  
Db 961 NDEEGDFAELYMEKWTLLIPPTTILINMGVVAGTSVAINSGYOSWGPLFGKLPFAFW 1020  
QY 1021 VIVHLXPFLLKGLMGRNRPFTIIVWAVLLASIFSLWLWRVDPFTTFLAGPNQTGCGINC 1080  
Db 1021 VIVHLXPFLLKGLMGRNRPFTIIVWAVLLASIFSLWLWRVDPFTTFLAGPNQTGCGINC 1080

RESULT 2

US-09-900-237-33  
; Sequence 33, Application US/099000237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13

; PRIOR APPLICATION NUMBER: 09/  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 33  
; LENGTH: 1065  
; TYPE: PRT  
; ORGANISM: Arabidopsis thaliana  
US-09-900-237-33

Query Match 78.0%;  
Best Local Similarity 78.4%;  
Matches 836; Conservative

QY 19 CQICADGLGTLTGDGVFTACDV  
Db 20 CQICSDNKGTVGDGRFVACDI  
QY 79 GEEGDDTDADGGS-DFNYPASG  
Db 80 GDNKDEGLADEGTEFNPY-  
QY 138 KYDSGEIPRGYVPSVTNSQ-MS  
Db 128 -----EYSHNHLPRLTSRQDTS  
QY 196 FSGSIGNVANKERVDGKMKQD  
Db 183 DPVGLGNVANKERVDGKMKQD  
QY 256 ETRQPLSRKVPFIASSKINPYRM  
Db 239 EAQPLSRKVSIFSSRINPYRM  
QY 316 ALSWILDQPKWFPINRETYLD  
Db 299 ALSWILDQPKWFPINRETYLD  
QY 376 VLSILAVDYPVDKVSICYSDDG  
Db 359 VLSILAVDYPVDKVSICYSDDG  
QY 436 QKIDYLDKDVQPSFVKDRRAK  
Db 419 AKIDYLDKDVQPSFVKDRRAK  
QY 496 RDPHGMIOVFLGHSGLDTEGN  
Db 479 GDHPGMIQVFLGONGGLDAEGN  
QY 556 QYMLNLDCHYINNSKALREAK  
Db 539 PFIILNLDCHYINNSKALREAK  
QY 616 DINLRGLDGIQGPVYVGTGCVF  
Db 599 DINLRGLDGIQGPVYVGTGCVF  
QY 674 SSDKKSNKHVDSSVPVFNLED  
Db 659 ESKDKKSGRHTDSTVPVFNLED  
QY 734 MEYGGVQPSQSTPESLKEAHLV  
Db 719 MENGVPVPSATPENLLKEAHLV  
QY 794 SVTCMKRPAFKGSAPINLSDR  
Db 779 SIYCMKPLPAFKGSAPINLSDR  
QY 854 INTTIYPLTSLPLLVYCIILPAI  
Db 839 VNTTIYPTITSIPLLMYCTLLAV

3  
; re 4507; DB 10; Length 1065;  
; id. No. 0;  
; Mismatches 107; Indels 26; Gaps 8;  
CRPCYBHERKEGTQACLOCKTKTKYRHRGSPAIR 78  
CRPCYBHERKEGTQACLOCKTKTKYRHRGSPAIR 79  
QKIDRMRSMRMTNGSGNVGHPKYDSGEIGLS 137  
EKISERMLGWHLTRGKGEEMGEPOYDK----- 127  
ASPDHMMSTGNISRRAPFPY-VNHSFNPSPRE 195  
ASPERLSVSTIAGGRKRLPYSSDVQNSPNRRIV 182  
MTNGTSTAPSEGRAATDIDASTEVNMDALLND 255  
V---STQAASE-RGGVDIDASTDILADEALLND 238  
LVVLSIFLHYRLTNPNVAYPLWLLSVCEIWF 315  
LVILCLFLHYRITNPVPNAPALVSVCEIWF 298  
YDREGEPQSLAAVDIFVSTVDPLKEPPIVANT 375  
YDREGEPQSLAAVDIFVSTVDPLKEPPIVANT 358  
FDALAEITSEFARKWVPVKKYDIEPRAPPEYFC 435  
FESLAETSEFARKWVPVKKYDIEPRAPPEYFA 418  
FKIRINALVSKALKCPBEQWMDGTWPNGNNT 495  
FKIRINALVSKALKCPBEQWMDGTWPNGNNT 478  
VYVSREKRPFGQHHKKAGANNALVRVSAVLNTG 555  
VYVSREKRPFGQHHKKAGANNALVRVSAVLNTG 538  
PNLGPVQYVQFPQRFQDIDRNDRYANRNTVFF 615  
PNLGPVQYVQFPQRFQDIDRNDRYANRNTVFF 598  
YGYEPPPIKA--KKPGFLASICGGKKKASKSKKR 673  
YGYEPPPIKA--KKPGFLASICGGKKKASKSKKR 658  
EGAGFDDKSKVLSMSQMSLEKRFQSAFVASTL 733  
EGAGFDDKSKVLSMSQMSLEKRFQSAFVASTL 718  
EDKSEWGTGWTGIVGVTEDILTGFKNHARGWR 793  
EDKSDMGMEIGWYGVSTEDILTGFKNHARGWR 778  
RWALGSVEILLFSRHCPLWYIGRGLKFLERFAY 853  
RWALGSVEILLFSRHCPLWYIGRGLKFLERFAY 838  
KFTMPISNLIASIFIALFISIFATGILEMRWS 913  
QFIIPQISNLIASIFIALFISIFATGILEMRWS 898



QY 914 GVGIDWNRNEQFWIGISAHFLFAVFOGLKVLVLAGIDTNTVTYSKANDSEGDFAEIYMF 973  
Db 899 GVGIDWNRNEQFWIGISAHFLFAVFOGLKVLVLAGIDTNTVTYSKANDSEGDFAEIYLF 958  
QY 974 KWTLLIPPTLLIINMVGWVAGTSYAINSGYQSWGPLFGKLFKFAFWIVHLYPFLKGLM 1033  
Db 959 KWTLLIPPTLLIINMVGWVAGTSYAINSGYQSWGPLFGKLFKFAFWIVHLYPFLKGLM 1018  
QY 1034 GRQNRPTIIVWVALLSIFSLWVRVDPPTTRLAGPNIOTCGINC 1080  
Db 1019 GRQNRPTIIVWVALLSIFSLWVRIDPFTSRVTPGDIILECGINC 1065  
RESULT 3  
US-09-838-539-8  
; Sequence 8, Application US/09838539  
; Patent No. US20020129401A1  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: 15621/03/US  
; CURRENT APPLICATION NUMBER: US/09/838,539  
; CURRENT FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/029,987  
; PRIOR FILING DATE: 1996-10-29  
; PRIOR APPLICATION NUMBER: 08/960,048  
; PRIOR FILING DATE: 1997-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 8  
; LENGTH: 881  
; TYPE: PRT  
; ORGANISM: Oryzae sativa  
US-09-838-539-8  
Query Match 77.1%; Score 4455.5; DB 10; Length 881;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 823; Conservative 39; Mismatches 18; Indels 1; Gaps 1;  
QY 201 GNVAKRVDGKMKODKAIPMTNGTSTIAPSEGRAATDIDASTYNNMEDALLNDETRQP 260  
Db 1 GNVAKRVDGKMKODKAIPMTNGTSTIAPSEGRAATDIDASTYNNMEDALLNDETRQP 60  
QY 261 LSRKVPASSKNIPYRMVTLRLVLSIFLHYRLTNPNRNAYPLWLLSVICEIWFALSWI 320  
Db 61 LSRKVPASSKNIPYRMVTLRLVLSIFLHYRLTNPNRNAYPLWLLSVICEIWFALSWL 120  
QY 321 LDQFPKWFPIINRETYLDRALRYDREGPSQLAAVDFVSTVDPLKEPPIVYANTVLSIL 380  
Db 121 LDQFPKWFPIINRETYLDRALRYDREGPSQLAAVDFVSTVDPLKEPPIVYANTVLSIL 180  
QY 381 AVDYPVDKVCYSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPEEYFCOKIDY 440  
Db 181 AVDYPVDKVCYSDDGASMLTFDALAETSEFARKWVPFVKYDIEPRAPEEYFSQIDY 240  
QY 441 LKDKVPSFKVKKORAMKREYEEFKIRINALYSKALKVPPEEGWIMQDGTWPFGNNTRDHPG 500  
Db 241 LKDKVPSFKVKKORAMKREYEEFKIRINGLVAKQKVPPEEGWIMQDGTWPFGNNTRDHPG 300  
QY 501 MIQVFLGHSGLDTEGNEPLPRVYVSREKRGFGQHHKAGAMNALVRYSAVLTNGQYMLN 560  
Db 301 MIQVFLGHSGLDTEGNEPLPRVYVSREKRGFGQHHKAGAMNALVRYSAVLTNGQYMLN 360  
QY 561 LDCDHYINNSKAVREAMCFMIDPDLGPOVCYVQFPQREDFGIDRNDRYANRNTVFEIDNLR 620  
Db 361 LDCDHYINNSKALREAMCFMIDPDLGPOVCYVQFPQREDFGIDRNDRYANRNTVFEIDNLR 420  
QY 621 GLDGIQGPVYVGTGVNFNTALYGYEPPPIKAKKPG-FIASLCCGKKKASKSKSSDKKK 679  
Db 421 GLDGIQGPVYVGTGVNFNTALYGYEPPPIKAKKPG-FIASLCCGKKKASKSKSSDKKK 480

QY 680 SNKHVDSSVPYFVNLEDIEEGVEGAGFDDKESVLMQSLMSEKRFQSQSAFVASTILMEYGGV 739  
Db 481 SNKHVDSSVPYFVNLEDIEEGVEGAGFDDKESVLMQSLMSEKRFQSQSAFVASTILMEYGGV 540  
QY 740 PQSSTPESLLKEAIVHISCGYEDKSEWGTGTEIGWYGVSTEDILTGFPMHARGWRSYVCMY 799  
Db 541 PQSSTPESLLKEAIVHISCGYEDKTEWGTGTEIGWYGVSTEDILTGFPMHARGWRSYVCMY 600  
QY 800 KRPAPKGSAPINLSDRLNQLRWALGSEVILFSRHCHPLWYGGRLKFLERFAYINTIY 859  
Db 601 KRPAPKGSAPINLSDRLNQLRWALGSEVILFSRHCHPIWYGGRLKFLERFAYINTIY 660  
QY 860 PLTSLPLLVYICILPAICLLTGKFIIMPETISNLASIFLALFISFATGILEMRWSGVGIDE 919  
Db 661 PLTSLPLLVYICILPAICLLTGKFIIMPETISNPFASIFLALFISFATGILEMRWSGVGIDE 720  
QY 920 WWRNEQFWVIGGISAHFLFAVFOGLKVLVLAGIDTNTVTYSKANDSEGDFAEIYFKWTTLL 979  
Db 721 WWRNEQFWVIGGISAHFLFAVFOGLKVLVLAGIDTNTVTYSKANDSEGDFAEIYFKWTTLL 780  
QY 980 IPPTTILINMVGWVAGTSYAINSGYQSWGPLFGKLFKFAFWIVHLYPFLKGLMGRQNR 1039  
Db 781 IPPTTILINMVGWVAGTSYAINSGYQSWGPLFGKLFKFAFWIVHLYPFLKGLMGRQNR 840  
QY 1040 PTIVIVWVALLSIFSLWVRVDPPTTRLAGPNIOTCGINC 1080  
Db 841 PTIVIVWVALLSIFSLWVRIDPFTSRVTPGDTQTCGINC 881  
RESULT 4  
US-09-900-237-26  
; Sequence 26, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: B81170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 26  
; LENGTH: 1091  
; TYPE: PRT  
; ORGANISM: Impatiens balsamia  
US-09-900-237-26  
Query Match 69.6%; Score 4021.5; DB 10; Length 1091;  
Best Local Similarity 68.1%; Pred. No. 0;  
Matches 749; Conservative 138; Mismatches 160; Indels 53; Gaps 20;  
QY 2 DGDADALASGRHAGDVQICADGLTDLGDFVTACDVCRFPVCRPCYERHERKEGTQAC 61  
Db 24 DSDGQPKPLKEANGQICQICGDTVGKTSATGDTFVACNECGFPVCRPCYERHERKDNQCC 83  
QY 62 LOCKYKXKRRHSGPAIRKEEGDDTDADD-GSDFNVPASGTEDQKQKTADMRSHRMWTGG 120  
Db 84 POCKTRYKQKSGPVEGDE-EEEDVDLENEFNY--SGKGKNOKKYTTARRPMQ----- 135  
QY 121 SCNVGHYPYDSGEIGLSKYVDSGEIPRGVYPSVTNS-QMSGEIPGASPDHMM-----SPTG 175  
Db 136 ---GDQDDIELSVSSSRHDESQRP---VPLLTGHSHSVSGEIP--TFDNHSIRTSPIG 186  
QY 176 NTSRRAPFPYVN-HSP-----NPSREF-SGSIGNVAKRVDGKMKODKAIPMTNG 226  
Db 187 PVKE--SIPYIDPQPAVRAIIVDPKSLNSYGLCNVDKERVSGWKLKQEKNNVQMTS- 243

QY 227 TSAPSEGRATDIDASTEYNMEDALLNDETROPLSRKVPFIASSKINPYRMVIVLRVVL 286  
Db 244 ---RYPEGKDTBTGSGNEELQMAA--DDIQPMRSRVPISSTHLPYRVVILRLIIL 298  
QY 287 SIFLHYRLNPNRYNAPLWLLSWICELWALSILWIDQFVKFPINREYTLDRALRYDRE 346  
Db 299 GFLQYRCHTHPVKDAVPLWLTSTICEVWFALSNLWLDQFVKFPVNRNREYTLDRLSMRFRE 358  
QY 347 GEPSOLAANDIFVSTVDPLKEPPIVNTANTVLSILAVDYPVKVSCVSDGASMLTFDAL 406  
Db 359 GEPSQLAPIDVFVSTVDPLKEPPIVNTANTVLSILAVDYPVKVSCVSDGASMLTFDAL 418  
QY 407 AETSEFARKWVPVKYIDTEPRAPRYFCOKIDYLDKQVQSFVDRRAMKREYEFKRI 466  
Db 419 SETAEFAKKWAPCKKHSSTPRAPRYFAOKIDYLDKQVQSFVDRRAMKREYEFKRI 478  
QY 467 INALVSKALKVPEEGWIMODGTWPVGNNTRDHPGMTQVFLHSGGLDTEGNEPLRLVYVS 526  
Db 479 INALVAKAQVPEEGWIMODGTWPVGNNSRDHPGMTQVFLHSGGLDTEGNEPLRLVYVS 538  
QY 527 REKRGFQHHKAGAMNALVRVSAVLNTNGQYMLNDCDHYINNSKAVREAMCFMDPNLG 586  
Db 539 REKRGFQHHKAGAMNALVRVSAVLNTNGAYLLNVDCDHYFNNSCKLKEAMCFMDPNLG 598  
QY 587 PQVYVQFQRFQDIDRNDRYANRNTVFFDINLRLDGIQGVYVGTGCVFNKRTIAYGE 646  
Db 599 KKTGYVQFQRFQDIDRNDRYANRNTVFFDINLRLDGIQGVYVGTGCVFNKRTIAYGE 658  
QY 647 PPKAK--KPG--FLASLCGGKKKASKKRSDDKSKNKHVDSVVPVFNLEIDRGEVGA 703  
Db 659 PVLTEDELFNIIKSCCSRSKGGKNGKYIDKNRALKRTSTAPIFNMEDIEGIE-- 716  
QY 704 GFDEKSVLMSQMSLEKRFQSAAFVASTMEYGGVQSQSTPESLLKEAHHVISCYEDK 763  
Db 717 GYDERSERLMAQ--SYEKRFQSPVLIAATFMEQGLPPSTNSATLLKEAHHVISCYEDK 775  
QY 764 SEWTEIGNYGVSVDITLTKFMHARGWRSVYCMKRPAPKGSAPINLSRLNQVLRWA 823  
Db 776 TENGKEIGNYGVSVDITLTKFMHTRGWISYCMPPKAPKGSAPINLSRLNQVLRWA 835  
QY 824 LGSVELFRHRPLWTGYGGRKLFLERFAYINTIYPTLSLPLLVYCIPLCLLTGKFI 883  
Db 836 LGSIELLSRHPCPWYGYSGRLKFLERLAYINTIYPTLSLPLLVYCIPLCLLTGKFI 895  
QY 884 MPEISNLASTWFTALFLSFATGILEMRWSGVGIDENWNEQFWIGGISAILFAVFOGL 943  
Db 896 VPESINYSATWFTLLFVSIFSGILELRWSGVTLDEWNEQFWIGGISAILFAVFOGL 955  
QY 944 LKVLADIDNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTILINMVGVAAGTSYAINS 1003  
Db 956 LKVLADIDNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTILINMVGIVAGVSFAINS 1015  
QY 1004 GYOSWGPFGKLEFFAFWVIVHLYPFLKGLMGQRNTPPTIIVVAVLLASIFSLWVRDP 1063  
Db 1016 GYOSWGPFGRLFFAFLWVIVHLYPFLKGLMGQRNTPPTIIVVAVLLASIFSLWVRDP 1075  
QY 1064 FT---TRLAGPNIQTGGINC 1080  
Db 1076 FTSDSTKARG---QCGIDC 1091

## RESULT 5

US-09-900-237-8  
; Sequence 8, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Syntheses  
; FILE REFERENCE: BB170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14

; PRIOR APPLICATION NUMBER: PCT/US98/015871

; PRIOR FILING DATE: 1999-07-13

; PRIOR APPLICATION NUMBER: 09/033

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: Microsoft Office 97

; SEQ ID NO 8

; LENGTH: 1165

; TYPE: PRT

; ORGANISM: zea mays

US-09-900-237-8

## Query Match

Best Local Similarity 66.58%

Matches 723; Conservative

Mismatches 175; Indels 69; Gaps 15;

9 KSGRHGAGDV-----CQICAE

96 ESGRAGGGAARRAPCOICG

63 QCKYKXKRRGSPAIRGEGDE

156 QCRTRYKRLKCPRVAGDEEE

122 GNVGHPKYDSCIEIGLSKYDSC

216 GD-AHP-----CQICAE

176 NISRRAPPYVYN-----HS

258 GSKRIHLPFPADPLVPQPS

228 SIAPSEGRAADTIDASTEYNM

310 QHVRSEGGDWDGDA-----

288 IFLHYRLNPNRYNAPLWLLS

364 FFHYRYVHMPKARDAFALWLS

348 EPSOLAAVDIEFVSTVDPLKEP

424 QPSQLAPIDFVSTVDPLKEP

408 ETSFARKWVPVKYIDTEPR

484 ETSFARKWVPVKYIDTEPR

468 NALYSKALKVPEEGWIMODGT

544 NALYAKAQVPEEGWIMODGT

528 EKRFGFQHHKAGAMNALVRV

604 EKRFGYNNHKKAGAMNALVRV

588 QVCYVQFPQRFQDIDRNDRYA

664 KVCYVQFPQRFQDIDRNDRYA

648 PIKAKP-----GFLA

724 P-KTKKPPSRTCNCPKWCLS

696 IEEGVGAGFQDDEKSVLMSQ

783 IDEGAPGA--DIEKAGIVNO

756 ISCYEDKSEWGTGIGWYGS

841 ISCYEDKTDGKBEIGWYGS

816 LNQVLRWALGSVEILFSRHC

/15871

33

are 3844.5; DB 10; Length 1165;

ed. No. 1.1e-317;

Mismatches 175; Indels 69; Gaps 15;

Db 901 LHOVLRLWAGSVEIFFSKCPLWYGYGGGLKFLERFSYNSIVPWTSLPILAYCTLPAL 960  
QY 876 CLLTGKFIMPEISNLASIFALFLSIFATGILEMRWSGVGIDEMWRNEQFQWVIGGISAH 935  
Db 961 CLLTGKFITPELTNVAISWFMALFICISVTGILEMRWSGVDAIDWWRNEQFQWVIGGSAH 1020  
QY 936 LFAVFOGLLVLKLAGIDTNTVTSKANDREGDPAELXMEKWTLLIPPTTILLIINWGVWA 995  
Db 1021 LFAVFOGLLVLKLAGIDTNTVTSKAGDDE-EFSELYTTFKWTLLIPPTTILLIINWGVWA 1079  
QY 996 GTSYAINSGYQSGPLFGKLFFAFWIVHLYPFLKGLMGRQNRPTTIVINWAVILLASIFS 1055  
Db 1080 GISNAINNGYESGPLFGKLFFAFWIVHLYPFLKGLVGRQNRPTTIVINWAVILLASIFS 1139  
QY 1056 LFWVRVDPPTTRLAGNIOTCGINC 1080  
Db 1140 LFWVRVDPFLAKSNGPLLEECGLDC 1164  
RESULT 6  
US-09-900-237-14  
; Sequence 14, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 14  
; LENGTH: 1039  
; TYPE: PRT  
; ORGANISM: Glycine max  
; NAME/KEY: UNSURE  
; LOCATION: (201)  
US-09-900-237-14  
Query Match 66.1%; Score 3820.5; DB 10; Length 1039;  
-Best Local Similarity 66.8%; Pred. No. 1e-315;  
Matches 719; Conservative 120; Mismatches 155; Indels 83; Gaps 15;  
QY 16 GDVCQICADGLTLDGVFTACDVCRCPCYEHKEGCTQACLOCKTKYKRHRGSP 75  
Db 34 GQVCEICGCVGLTVBDGLFVACNECGFPVCRPCYERREGSHLCPOCKTRYKRLKSP 93  
QY 76 AIRCEEGDDTDAD-GSDENYPASGTEDOKOK--IADRMRSRMNTGSGNVGHPKYDS 131  
Db 94 RVEGDD-DEEDVDIEHEFN----IDEQKNHGQVAEAMLHGRMSYERG-----PEDD- 141  
QY 132 GEIGLSKYDSGEIPRGVPSVTNSQMSGRIPIGASPDHMMSPGTGNISRRAPPPVYVNHSPN 191  
Db 142 -----DNSQFTPTPIAGGRSPVSGEPFPISSNAYGQDMLSSSLHKKRV-HPYVPSBPG 192  
QY 192 PSR---EFGSGIGNAWKERVGVKMKODKGAIPMTNGTSTAPSEGRAAATDIDASTEYN 247  
Db 193 SARWDEKXDG-----WKDRMDWKLQ-----GNLGEPEPDED----- 225  
QY 248 MEDALLNDETQPLSRKVPITASSKINPYRMVIVLRVLWLSIFLHYLRNLNPNVRYNAYPLWLL 307  
Db 226 -PDAAMLDEARQPLSRKVPITASSKINPYRMVIVLRVLWLSIFLHYLRNLNPNVRYNAYPLWLL 284  
QY 308 SVICEIWFALSWILDOPPKWFPINRETYLDRLALRYDRGEPSSQLAAVDIVFVSTVDPLKE 367

Db 285 STICEIWFALSWILDOPPKWFPIDRETYLDRLALRYDRGEPSSQLAAVDIVFVSTVDPLKE 344  
QY 368 PPVITANTVLSLAVDVPVKVSCYSDGASMLTFDAAETSEFARKWVPFYKYYDIBP 427  
Db 345 PPLVNTANTVLSLAMDVPDKISCIYISDDGASMTFESLSETAEFARKWVPFYKYYDIBP 404  
QY 428 RAPEFYFCQKTDYLDKQVPSFVDRRAMKREYEEFKIRINALVSKALKVPBEGWLMQDG 487  
Db 405 RAPEMYFSEKTDYLDKQVPTFVARRAMKREYEEFKIRINALVSKALKVPBEGWLMQDG 464  
QY 488 TPWGNNTTRDHPGMIQVFLGSHSGGLDTEGNELPLVYVVSREKRPFGHKKAGAMNALVR 547  
Db 465 TPWGNNTTRDHPGMIQVFLGSHSGGLDTEGNELPLVYVVSREKRPFGHKKAGAMNALVR 524  
QY 548 VSAVLNTNGYMLNLDCHYINNSKAVREAMCFMLDPLGPOVCYVQFPORFDCIDRNDY 607  
Db 525 VSAVLNTNAPFMLNLDCHYVANSKAAEAMCFMLDPLGPOVCYVQFPORFDCIDRNDY 584  
QY 608 ANRNTVFFDINLRGLDGTQGPVYVGTGVFNRTAIYGYEPTIKAKRPGFLA---SICGGK 664  
Db 585 ANRNTVFFDINMKGLDGTQGPVYVGTGVFNRTAIYGYEPTIKAKRPGFLA---SICGGK 644  
QY 665 KKASKSKRSSDKKSKKHVDSSVPFNLEDIEBEGVAGAGFDDKSVLMSQMSLEKRFQ 724  
Db 645 RKYKENDANGAASLK-----GMDDDKEVLMSQMFKKFQ 683  
QY 725 SAAVASTLMEYGGVQPSSTPESLLKEAHVISCYEDKSEWTEIGWYGVSTEDILTG 784  
Db 684 SSIFVTSTIMESGGVPPSSPAALLKEAHVISCYEDKSEWTEIGWYGVSTEDILTG 743  
QY 785 FKMHWGRWSYVCMKPAKFGAPINLSRLNOVLWALGSVEILFSRHCPLWYGY-GG 843  
Db 744 FKMHWGRWSYVCMKPAKFGAPINLSRLNOVLWALGSVEILFSRHCPLWYGY-GG 803  
QY 844 RLKFLERFAYINTTITPLTSLPLVYVLCILPAICLLTGKFIPEISNLASIFALFLSIF 903  
Db 804 KLKWLERFAYANTTIVYPTSLPLVYVLCILPAICLLTGKFIPEISNLASIFALFLSIF 863  
QY 904 ATGILEMRWSGVGIDEMWRNEQFVIGGISAHLFAVQGLLKVLAGIDTNTVTSKANDE 963  
Db 864 ATGILEMRWSGVGIDEMWRNEQFVIGGISAHLFAVQGLLKVLAGIDTNTVTSKANDE 923  
QY 964 EGDFAELYNFKWTTLLIPPTTILLIINWGVVAGTSYAINSGYQSGPLFGKLFFAFWIV 1023  
Db 924 E-EFGELYTFKWTLLIPPTTILLIINWGVVAGTSYAINSGYQSGPLFGKLFFAFWIV 982  
QY 1024 HLYPFLKGLMGRQNRPTTIVINWAVILLASIFSLWVRVDPPTTRLAGNIOTCGINC 1080  
Db 983 HLYPFLKGLMGRQNRPTTIVINWAVILLASIFSLWVRVDPPTTRLAGNIOTCGINC 1039  
RESULT 7  
US-09-900-237-10  
; Sequence 10, Application US/09900237  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cellulose Synthases  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: US/09/900,237  
; CURRENT FILING DATE: 2001-07-06  
; PRIOR APPLICATION NUMBER: 60/092,844  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT/US99/15871  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/720383  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 10  
; LENGTH: 1086  
; TYPE: PRT  
; ORGANISM: Zea mays

US-09-900-237-10

Query Match 66.1%; Score 3818; DB 10; Length 1086;  
Best Local Similarity 64.7%; Pred. No. 1.8e-315;  
Matches 716; Conservative 141; Mismatches 178; Indels 72; Gaps 17;

QY 2 DGDADALKSRHGGAGDVQICADGLGTTLDGVDFTACDYCRFPVPCYEHKRGQTQAC 61  
DB 23 DGDGPP-KPPREQNGQVQICGDDVGLAPGDPFVACNECAFVPCVDCYERREGTQNC 81

QY 62 LQCKTKYKRRHSGPAIRREGDDTDADGSDENYPASGTEDQOKIADPWRWRMTGGS 121  
DB 82 PCKTRYKRLKUGCQRTGDEEDGVVDLNEFNWDG----HDSQSVAESML----- 128

QY 122 GNYCHPKYDSGEICLSKYDSGEIPRGY-----VPSVTNSQMSGEIPGASPDHMHMSPT-- 174  
DB 129 --YCHMSYGRGG-----DPNGAPAQFQLPNPNVPLLTNGQVDDIP---PEQHALVPSPM 177

QY 175 -GNISRAPPPYVN-----HSPNPSREFSG-SIGNVAKERVGDGKMKQDKGAIPMTN 225  
DB 178 GGGGKRHLPLFYADPSLPVQPSRMDPSKOLAAVYGSVAKKEMENMKQQRER--MHQTG 235

QY 226 GTSIAPSEGRAATDIDASTEYNNMEDALNDETROPKSRVPYIASSKINPYRMVIVLRLVY 285  
DB 236 -----NDGGDDGDDA-----DLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVV 282

QY 286 LSIFLHVLTLNPNVRYNAYPLWLLSVCEIWFALSWMILDOFPKWPFPINRETYLDRLALRYDR 345  
DB 283 LGFFFRHVMHPVNDFAFWLWSVCEIWFAMSWILDOFPKWPFPINRETYLDRLSLRFDK 342

QY 346 EGPFSQALAAVDIFSVTVDPKLEPIVANTVLSILAVDYPVKVSCYVSDGASMLTFDA 405  
DB 343 EGQPSQALPIDEFVSTVDPKLEPLVTNTVLSILSVDPVKVSCYVSDGAAMLTFEA 402

QY 406 LAETSEFAKWWPVFKYKDYDEPAPPEFYCOKIDYLDKDVQSPFYKVDRAKREYEEFKI 465  
DB 403 LSETSEFAKWWPVFKYKDYDEPAPPEFYCOKIDYLDKDVQSPFYKVDRAKREYEEFKI 462

QY 466 RINALYSKALKVPEEGWIMODGTPWPGNNTDRHPGMIQVFLGHSGGLDTEGNEPLRLVYV 525  
DB 463 RINALVAKAKVPEEGWIMODGTPWPGNNVDRHPGMIQVFLGHSGGLDTEGNEPLRLVYV 522

QY 526 SRKRPQGHKHKAGAMNALRVYSALVTNGOYMLNDCDCHYINNSKAVREACFMMDPNL 585  
DB 523 SRKRPQGHKHKAGAMNALRVYSALVTNAPYLLNDCDCHYINNSKAVREACFMMDPNL 582

QY 586 GPQVCYVQFPQREDGIDRNDRYANRNTVFFDINLRGLDIOGVPVYVGTGCVFNRTAIYGY 645  
DB 583 GKVCYVQFPQREDGIDRNDRYANRNVFFDINMKGLDIOGVPVYVGTGCVFNRTAIYGY 642

QY 646 EPPFKAKKP-----GFLASLGGKKKASKKRSDDKK--SNKHVDSVVPVFNL 693  
DB 643 DAP-KTKKPPSRCTNCWPKWCFCCCGFNKQKKTTPKTEKKLLFFKKEENQSPAYAL 701

QY 694 EDIEEGVAGFDEKSVLMSQMSLEKRFQSSAAFAVASTIMEYGGVPOSSSTESLLEKEAI 753  
DB 702 GEIDEAAPGA--ENKAGIVNQKLEKKFGQSSVFVTSTLLENGGLTKSASPASLLKEAI 759

QY 754 HVISCGYEDKSENGTEIGWYIGSVTEIDLTGFKMHARGWRSVYCMKPAKFAFGSAPINLS 813  
DB 760 HVISCGYEDKTDGKTEIGWYIGSVTEIDLTGFKMHARGWRSVYCMKPAKFAFGSAPINLS 819

QY 814 DRINQVLRWALGVSIFLFRHCPWTYGGYGRFLERFAYINTIYPLTSLPLLVYCIPL 873  
DB 820 DRHQVLRWALGVSIFLFRHCPWTYGGYGRFLERFAYINTIYPLTSLPLLVYCIPL 879

QY 874 AICLLTKGFTMPEISNLASTWFTALFLSPATGILEMRSGVGIDEMRWNEQFWTGGTGS 933  
DB 880 AICLLTKGFTMPEISNLASTWFTALFLSPATGILEMRSGVGIDEMRWNEQFWTGGTGS 939

QY 934 AHLFAVQGLLKVLAGIDTNTVTSKANDEGDFAEYKFKWTTLLIPTTILINNAVY 993  
DB 940 SHLFAVQGLLKVLAGIDTNTVTSKANDEGDFAEYKFKWTTLLIPTTILINNAVY 998

QY 994 VAGTSYAINSGYQSWGPLFGKLVIVHLYPFLKGLMGQRNRTPTIVVAVLLASI 1053  
DB 999 VAGYSNAINNGYESNGPLFGKLVIVHLYPFLKGLVGRNRTPTIVIVWSILLASI 1058

QY 1054 FSLWVRVDPPTTRLAGNIOPT 1080  
DB 1059 FSLWVRIDPFLAKDDGPLEET 1085

RESULT 8  
US-09-900-237-4  
; Sequence 4, Application US/099  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cell  
; FILE REFERENCE: Bb170 US CI  
; CURRENT APPLICATION NUMBER: 09/900,237  
; PRIOR APPLICATION NUMBER: 60/344  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PC  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: 09/83  
; PRIOR FILING DATE: 2000-12-2  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 9  
; SEQ ID NO 4  
; LENGTH: 1148  
; TYPE: PRT  
; ORGANISM: Zea mays  
US-09-900-237-4

Query Match 66.1%; Score 3818; DB 10; Length 1148;  
Best Local Similarity 65.1%; Pred. No. 2e-315;  
Matches 714; Conservative 141; Mismatches 168; Indels 74; Gaps 16;

QY 16 GDVQICADGLGTTLDGVDFTACDYCRFPVPCYEHKRGQTQACDYCRFPVPCYEHKRGQTQAC 75  
DB 95 GQVCQICGDDVGRNPDGEPFVLAFFICRDCYERREGTQNCPOCKTRFKRLKGA 154

QY 76 AIRGEGDDTDADGSDENYPASGTEDQOKIADPWRWRMTGGSVNVG-----HPKY 129  
DB 155 RVPGEDEGVDLNEFNWDG----HDSQSVAESMLHAHMSYGRGADLDGVPQFHP-- 209

QY 130 DSGEIGLSKYDSGEIPRGY-----VPSVTNSQMSGEIPGASPDHMHMSPT-- 186  
DB 210 -----HSPNPSREFSG-SIGNVAKERVGDGKMKQDKGAIPMTN 250

QY 187 N-----HSPNPSREFSG-SIGNVAKERVGDGKMKQDKGAIPMTN 238  
DB 251 DPNLVQPSRMDPSKOLAAVYGSVAKKEMENMKQQRER--MHQTG 302

QY 239 DIDASTEYNNMEDALNDETROPKSRVPYIASSKINPYRMVIVLRLVLSIFLHVLRLNVP 298  
DB 303 GDDA-----DLPLMDEARQOLSRKIPLPSSQINPYRMIIIRLVLCFFHFVYRMHPV 355

QY 299 RNAYPLWLLSVCEIWFALSWMILDOFPKWPFPINRETYLDRLALRYDR 358  
DB 356 PDFAFWLWSVCEIWFAMSWILDOFPKWPFPINRETYLDRLSLRFDKEGHFSQAPVDF 415

QY 359 VSTVDPKLEPIVANTVLSILAVDYPVKVSCYVSDGASMLTFDA 418  
DB 416 VSTVDPKLEPIVANTVLSILAVDYPVKVSCYVSDGASMLTFDA 475

QY 419 FVKYKDYDEPAPPEFYCOKIDYLDKDVQSPFYKVDRAKREYEEFKI 478  
DB 476 FCKRYKDYDEPAPPEFYCOKIDYLDKDVQSPFYKVDRAKREYEEFKI 535

QY 479 EPGWTMQDGTWPGNNTDRHPGMIQVFLGHSGGLDTEGNEPLRLVYV 538  
DB 536 EPGWTMQDGTWPGNNTDRHPGMIQVFLGHSGGLDTEGNEPLRLVYV 595

QY	539	AGAMNALVRVSAVLTNQYMLNLDCHYINNISKAVREAMCFMLDNPJGQVYVQFPQRF	539
Db	596	AGAMNALVRVSAVLTNAPYLLNLDCHYINNISKATKEAMCFMMDPLLGKKYCVYQFPQRF	655
QY	599	DGIDRDRYANRVTFEDINLRGLDIOGPVYVGGCVFNRTAIYGVPEPPIKAKKP----	654
Db	656	DGIDRDRYANRVTFEDINMKGLDIOGPYVGTGCVFRQALYGVDAKPYTKKPSPT	714
QY	655	-----GFIASLCGGKKKASKSKRSKD-----KSKNHVDSSVPVFNLEDIEBEGAG	704
Db	715	CNCWPKWICCCCFGNRKTTKTKTKPKFEKIKLKKKKENQAPAYALGEIDEAAPGA-	773
QY	705	FDEKSVLMSQSLKRFQGSAAFASTIMEYGGVPOSTPESSLKKAHIVISGYEDKS	764
Db	774	-ENEKASIVNQQLKKFKGQSVFVASTULLENGGTLKASAPASLLKKAHIVISGYEDKT	832
QY	765	EWTEIGTWIGSVTEDIITLTFKMHARGWSVCMKRPAPKGSAPINLSRLNOVLRAWL	824
Db	833	GWGKDIGTWIGSVTEDIITLTFKMHGCHGWSYICYPKRAAFKGSAPNLSRLHQVLRAWL	892
QY	825	GSVEILFSRHCPLWYGGYGRKLFERFAVINTIYPTLSPLLYCYCIPATCLLTGKFM	884
Db	893	GSTEIEFSNHCPLWYGGYGRKLFERFSYINSIYPTLSPLLYCYCIPATCLLTGKFI	952
QY	885	PETSNLASIWFALFISFATGILEMRWSGVGIDEMWRNEQFVWIGGISAHLFAVFGLL	944
Db	953	PELNNTASLWFLSFLCIFATSIEMRWSGVGIDEMWRNEQFVWIGVSSHFAVFGLL	1012
QY	945	KVLAGIDTNTVTSKANDEGQFAELYMFKWTTLLIPTTILINMWGVVAGTSYATNSG	1004
Db	1013	KVLAGVDTSTVTSKGGDE-BSLEYLYFKWTTLLIPTTILINMWGVVAGTSYATNSG	1071
QY	1005	YQSGWELFKLFPFAFWIVHLIYPTLKGMLGRQNRPTTIIVAVLLASIFSLMVRVDPF	1064
Db	1072	YESWGELFKLFPFAFWIVHLIYPTLKGMLGRQNRPTTIIVYMSILLASIFSLMVRIDPF	1131
QY	1065	TTRLAGPNIQTCCIG 1080	
Db	1132	LAKDDGPLLEECGLDC 1147	
RESULT 9			
US-09-838-539-6			
; Sequence 6, Application US/09838539			
; Patent No. US20020129401A1			
; GENERAL INFORMATION:			
; APPLICANT: Stalker, D. et al.			
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter			
; TITLE OF INVENTION: Sequences			
; FILE REFERENCE: 15621/03/US			
; CURRENT APPLICATION NUMBER: US/09/838,539			
; CURRENT FILING DATE: 2001-04-18			
; PRIOR APPLICATION NUMBER: 60/029, 987			
; PRIOR FILING DATE: 1996-10-29			
; PRIOR APPLICATION NUMBER: 08/960, 048			
; PRIOR FILING DATE: 1997-10-29			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 6			
; LENGTH: 974			
; TYPE: PRT			
; ORGANISM: Gossypim hirsutum			
; US-09-838-539-6			

Qy	78	RGE	DDT	DAD	GD	SDF	NYP	PAS	GTD	QK	IAD	RMR	SRW	MNT	GG	SNV	-	HP	KYD	S	E	I	G	L	136	
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:		
Db	59	----	D	E	N	L	T	D	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----		
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:		
Qy	137	SKY	S	G	E	I	P	R	G	Y	P	S	V	T	N	S	Q	M	S	G	E	I	P	C	A	196
Db	97	STL	D	S	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----		
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:		
Qy	197	SGS	I	N	N	V	A	M	K	E	R	V	D	G	M	K	O	K	G	A	I	P	M	T	N	256
Db	104	AED	N	G	N	S	I	W	K	N	R	V	E	S	W	K	E	K	N	K	K	P	A	T	----	154
Qy	257	TRO	P	L	S	R	K	V	P	A	S	S	K	I	N	P	Y	V	L	R	V	L	S	I	F	316
Db	155	S	-	Q	P	L	S	T	I	I	P	I	P	K	S	R	A	P	Y	F	T	V	I	M	R	213
Qy	317	LSW	I	D	O	F	P	K	W	P	I	N	R	E	T	Y	L	D	R	I	A	L	R	V	D	376
Db	214	FSW	I	D	O	F	P	K	W	P	V	N	R	E	T	Y	I	D	R	L	S	A	R	Y	E	273
Qy	377	LSI	L	A	V	D	P	Y	D	K	S	C	Y	S	D	G	A	S	M	L	F	D	A	L	E	436
Db	274	LSI	L	A	V	D	P	Y	D	K	S	C	Y	S	D	G	A	S	M	L	F	D	A	L	E	333
Qy	437	KID	L	K	D	K	V	O	P	S	F	K	R	R	A	M	K	E	Y	E	B	F	K	I	R	496
Db	334	KID	L	K	D	K	V	O	P	S	F	K	R	R	A	M	K	R	O	Y	E	E	K	I	R	393
Qy	497	DHP	C	M	I	O	V	F	L	G	H	S	G	L	D	T	E	G	N	E	L	P	R	L	V	556
Db	394	DHP	C	M	I	O	V	F	L	G	H	S	G	A	R	D	E	N	E	L	P	R	L	V	S	453
Qy	557	YML	N	D	C	D	H	I	N	N	S	K	A	V	R	E	A	M	C	F	L	M	D	N	L	616
Db	454	FIL	N	D	C	D	H	I	N	N	S	K	A	V	R	E	A	M	C	F	L	M	D	N	L	513
Qy	617	INL	R	G	L	D	I	O	G	P	V	T	G	C	Y	N	F	T	A	I	G	Y	B	P	P	676
Db	514	VNM	G	L	D	I	O	G	P	V	T	G	C	Y	N	F	T	A	I	G	Y	B	P	P	S	573
Qy	677	KKS	N	K	H	V	D	S	V	P	N	F	N	L	E	D	I	E	E	G	V	E	G	A	F	735
Db	574	LYR	A	K	R	E	E	L	D	A	A	I	N	F	U	R	E	I	D	----	----	----	----	----	----	627
Qy	736	YGG	V	P	O	S	T	P	E	S	L	L	K	E	A	H	I	V	S	C	G	E	D	K	S	795
Db	628	NGV	A	E	S	A	N	P	T	L	I	K	E	A	H	I	V	S	C	G	E	K	T	A	M	687
Qy	796	YCM	P	R	P	A	F	G	S	A	P	I	N	S	D	R	L	N	O	V	R	W	A	L	S	854
Db	688	YCM	P	R	P	A	F	G																		

**Qy**      18 VCQCADGLTTLGDVFTACDVCRFPPVCRPCYEHKKEGTQACLCKTKYKRHRGSPAI 77  
          || | : || : || | || : || : || : || : || : || : || : || : || : || : || :  
**Dd**      8 VCHTCGEHVGLNVNGEPFVACHECNFPICKSCFEYDLKEGRKKACLCR-----GSFY- 58

```

; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 32
; LENGTH: 701
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-900-237-32
```

```

Query Match      58.98; Score 3404.5; DB 10; Length 701;
Best Local Similarity 88.28; Pred. No. 1.2e-280;
Matches 618; Conservative 56; Mismatches 24; Indels 3; Gaps 2;
```

```

QY 383 DYPVDKVCYVSDGASMLTDLAETSEFARKWVPFVKYKIDIEPRAPFEFFCQKIDYK 442
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      1 DYPEKVCYVSDGASMLTDLAETSEFARKWVPFVKYKIDIEPRAPFEFFCQKIDYK 60

QY 443 DKVQSFVKDRAMKREYEEFKIRINALVSKALVPEEGWTMDGTWPNGNTRDHPGMI 502
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      61 DKVQSFVKERRAMKREYEEFKIRVNGLVAKAQVPEEGWTMDGTWPNGNTRDHPGMI 120

QY 503 QVFLGSGGLDTEGNEPRLVYVSRKRPQHHKAGAMNALVRVSAVLNFGQVMLND 562
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      121 QVFLGSGGLDTEGNEPRLVYVSRKRPQHHKAGAMNALVRVSAVLNFGAFLNLD 180

QY 563 CDHYINNSKAVREAMCFMDNLPQVCYVQFPQRFQDIDRNDRYANRNTVFEDINLRGL 622
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      181 CDHYINNSKALREAMCFMDNLPQVCYVQFPQRFQDIDRNDRYANRNTVFEDINLRGL 240

QY 623 DGIOGPVYVGTGCVFNRTAIGYEPPIKAK--KPGFLASLCGG-KKKASKSKRSSDKKK 679
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      241 DGIOGPVYVGTGCVFNRTAIGYEPPIKAK--KPGFLASLCGG-KKKASKSKRSSDKKK 300

QY 680 SNKHVDSVPVFNLEDIEEGVEGAGFDDKSKVLSMSQMSLEKRFQGSAAFAVASTLMYGV 739
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      301 SGKHVDSVPVFNLEDIEEGVEGAGFDDKSKVLSMSQMSLEKRFQGSAAFAVASTLMYGV 360

QY 740 PQSSTPESILLKEAIIHVISCGYEDKSEWCTEIGWIYGSVTEDILTCFKMHARGWRSYVCM 799
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      361 PQSATPETLLKEAIIHVISCGYEDKSEWCTEIGWIYGSVTEDILTCFKMHARGWRSYVCM 420

QY 800 KRPAFGSAPINLSRLNQVRLWALGSVEILFSRHCPLWYGYGGRKLEFARFAYINTIY 859
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      421 KRPAFGSAPINLSRLNQVRLWALGSVEILFSRHCPLWYGYGGRKLEFARFAYINTIY 480

QY 860 PLTSLPLLVCYLPAICLLTKGFMPEISNLASINWFIALFISIPATGILEMRWSGVGIDE 919
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      481 PVTATPLLWCTLPVAVCLLTNKFIIPOISNLASINWFIALFISIPATGILEMRWSGVGIDQ 540

QY 920 WRNRQFVWVIGGISAHFAVFGGLKVLGAGIDTNTFTVTSKANDDEGDFAEILMFKWTLL 979
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      541 WRNRQFVWVIGGISAHFAVFGGLKVLGAGIDTNTFTVTSKANDDEGDFAEILMFKWTLL 600

QY 980 IPTTILLIINMVGVSAGTSYAINSGYQSGWGLFGKLFPAFVVIHLYPFLKGLMGRQNR 1039
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      601 IPTTILLIINMVGVSAGTSYAINSGYQSGWGLFGKLFPAFVVIHLYPFLKGLMGRQNR 660

QY 1040 PTVIVWAVLLASIFSLWVRDPTTTLRAGPNITQTCGNC 1080
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      661 PTVIVWAVLLASIFSLWVRDPTTTLRAGPNITQTCGNC 701
```

```

RESULT 11
US-09-900-237-18
; Sequence 18, Application US/09/900,237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: 09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Glycine max
US-09-900-237-18

Query Match      56.04; Score 3234; DB 10; Length 793;
Best Local Similarity 73.38; Pred. No. 4.4e-266;
Matches 581; Conservative 87; Mismatches 87; Indels 14; Gaps 5;

QY 296 NPVRNAYPLMLLSVICIWEFAN 355
      :||: ||| ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
      4 HPVNDAYGLMLTSLVICIWEFAN 63

QY 356 DIFVSTVDPLKEPPIVNTANTV 415
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      64 DVFVSTVDPMKEPLIANTV 123

QY 416 WVPVKYKIDIEPRAPFEFFCQKIDYK 475
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      124 WVPCKKYNIEPRAPFEFFCQKIDYK 183

QY 476 KVPEEGTMDGTWPNGNTR 535
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      184 KVPEEGTMDGTWPNGNTR 243

QY 536 HKKAGAMNALVRVSAVLNFGQVMLND 595
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      244 HKKAGAMNALVRASIIITNAP 303

QY 596 QRFQDIDRNDRYANRNTVFEDINLRGL 655
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      304 QRFQDIDRNDRYANRNTVFEDINLRGL 363

QY 656 -----FLASLCGGKKKAS 707
      : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
      364 KTCNCWPKWCCLCCGSRKKKY 419

QY 708 EKSVMLSQMSLEKRFQGSAAFAVASTLMYGV 767
      |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||
      420 EKTSLNLTQTKLEKRFQGSAAFAVASTLMYGV 479

QY 768 TEIGWIYGSVTEDILTCFKMHARGWRSYVCM 827
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      480 KEVGIYGSVTEDILTCFKMHARGWRSYVCM 539

QY 828 EILFSRHCPLWYGYGGRKLEFARFAYINTIY 887
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      540 EILFSRHCPLWYGYGGRKLEFARFAYINTIY 599

QY 888 SNLASINWFIALFISIPATGILEMRWSGVGIDE 947
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
      600 SNLASINWFIALFISIPATGILEMRWSGVGIDE 659

QY 948 AGIDTNTFTVTSKANDDEGDFAEILMFKWTLLIPTTILLIINMVGVSAGTSYAINSGYQSGWGLFGKLFPAFVVIHLYPFLKGLMGRQNR 1007
```

```
Db 660 AGVNTFTVTSKAAD-DGEFSELYIFKWTSLIPMTLLIMNIVGVVIGSADINNGYDS 718
QY 1008 WGPLFGKLFPAFWVIVHLVPLFKGLMGRNRPFTIIVWAVLASISFISLLWVRVDPFTTR 1067
Db 719 WGPLFGKLFPAFWVIVHLVPLFKGLMGRNRPFTIILVWSILLASILLTLMWVRINPFVSR 778
QY 1068 LAGPNIQTCGINC 1080
Db 779 -DGPVLEICGLNC 790

RESULT 12
US-09-900-237-24
; Sequence 24, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 24
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Picramnia pentandra
US-09-900-237-24

Query Match 52.6%; Score 3038.5; DB 10; Length 740;
Best Local Similarity 74.1%; Pred. No. 1.6e-249;
Matches 552; Conservative 97; Mismatches 79; Indels 17; Gaps 6;

QY 345 REGEPSOLAVIDFVSTVDPLKEPPDIVTANTVLSILAVDPYDKVSCYVSDGASMLTFD 404
Db 3 KEKPSLGLDIFVSTVDPMKEPPDIVTANTVLSILAVDPYDKVTCYVSDGGAAMLTFE 62
QY 405 ALAETSEFARKWPFVVKVYDIEPRAPEFYFCOKIDYLDKVKOPSFVKDRRAMKREYEEPK 464
Db 63 ALSETSEFARKWPFVVKVYDIEPRAPEFYFCOKIDYLDKVKOPSFVKDRRAMKREYEEPK 122
QY 465 IRINALVSKALKVPEEGWIMQDGTWPNGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVY 524
Db 123 VRINGLVAMAQKVPEDGWTMQDGTWPNGNVRDHPGMIQVFLGHNGVRDVEGNEPLRLY 182
QY 525 VSREKRPGFQHHKAGAMNALVRSAVLNGOYMLNDCDHYINNSKAVREACFLMDPN 584
Db 183 VSREKRPGFQHHKAGAMNSLVRSAVINAPYILNVDCDHYINNSKALREACFLMDPT 242
QY 585 LGPOVCYVQPPQRFQDGRNDRYANRNTVFFDINRLGLDGIQGPVYVGTGCVFNRTAIVG 644
Db 243 SGKKLCYVQPPQRFQDGRNDRYANRNVFFDINMKGLDGIQGPVYVGTGCVFRRVLYG 302
QY 645 YEPIKAKKPG-----FLSLCGGKK-KASKSKRSDDKSKNKHVDSSVPFNLED 695
Db 303 YDAPVTTKSGKACNCPWKLCCCGGSRKSKKSKKPKK-----EKKSKNREASQKHALEN 358
QY 696 IEEGVEGAGFDEKSVLMSQMSLEKRFQGSAAFAVASTLMEYGGVQSGSTPESLLKEAIVH 755
Db 359 IEEGM--GGLNSEKSCETTPLEKKEGQSPVFAVASTLLEDGGVPDAPALKEAIVQ 416
QY 756 ISCYEDKSEWGTIEGNIYGSVTEIDILTGFKMHARGWRSVYCMKPAFKGSAPINLSDR 815
Db 417 ISCYEDKTEWGEVGIYGSVTEIDILTGFKMHARGWRSVYCMKPAFKGSAPINLSDR 476
QY 816 LNOVLRWALGSVELFSRHCPLWYGGRLKFLERFAYINTTIYPLTSLPLLYVCTLPAL 875
```

```
Db 477 LHQVLRWALGSVELFSRHCPLWYGGRLKFLERFSYSSVYVWTSIPLLYVCTLPAL 536
QY 876 CLITGKREIMEISINLASINFIATGILEMWSGVGIDEMWRNEQFVWIGGISAH 935
Db 537 CLITGKREIVEISINLASINFIATGILEMWSGVGIDEMWRNEQFVWIGVSSH 596
QY 936 LEAVFQGLLKVLAGIDTNTFTVTSKANDGEGFAELMFKWTLLIPPTTILINMVGVYA 995
Db 597 LEALFQGLLKVLAGVNTFTVTSKAAD-EGDFSELYLFKWTLLIPPTTILINIVGVV 655
QY 996 GTSYAINSGYOSQSGPLFGKLFPAFWVIVHLVPLFKGLMGRNRPFTIIVWAVLASIFS 1055
Db 656 GVSDAINNGYDSWGPLFGRFFAFWVIVHLVPLFKGLLQKQDRTPTIIVWSILLASILT 715
QY 1056 LLWVRVDPFTTRLAGPNIQTCGINC 1080
Db 716 LLWVRINPFVSR-DGPVLEVCGLNC 739

RESULT 13
US-09-900-237-31
; Sequence 31, Application US/09900237
; Patent No. US20020120124A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen
; TITLE OF INVENTION: Plant Cellulose Synthases
; FILE REFERENCE: BB1170 US CIP
; CURRENT APPLICATION NUMBER: US/09/900,237
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: 60/092,844
; PRIOR FILING DATE: 1998-07-14
; PRIOR APPLICATION NUMBER: PCT/US99/15871
; PRIOR FILING DATE: 1999-07-13
; PRIOR APPLICATION NUMBER: 09/720383
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Gossypium hirsutum
US-09-900-237-31

Query Match 48.7%; Score 2816; DB 10; Length 685;
Best Local Similarity 74.0%; Pred. No. 1.1e-230;
Matches 509; Conservative 78; Mismatches 77; Indels 24; Gaps 6;

QY 414 RKWVPVFKVYDIEPRAPEFYFCOKIDYLDKVKOPSFVKDRRAMKREYEEFKIRINALVSK 473
Db 1 RRWVPFCKKHVPEPRAPEFYFNEKIDYLDKVKHPSFVKERRAMKREYEEFKIRINALVAK 60
QY 474 ALKVPESGWLMDGTPWPGNTRDHPGMIQVFLGHSGGLDTEGNEPLRLVYVSRKRPGF 533
Db 61 AOKPESGWNQDGTWPNGNTRDHPGMIQVILGSAGALVDGKELPRLVYVSRKRPGY 120
QY 534 QHHKKAGAMNALVRSAVLNGOYMLNDCDHYINNSKAVREACFLMDPNLGPQVCYVQ 593
Db 121 QHHKKAGAENALVRSAVLNAPFLNLDCHYINNSKAREACFLMDPQFGKLCYVQ 180
QY 594 FPQRFQDGRNDRYANRNTVFFDINRLGLDGIQGPVYVGTGCVFNRTAIVGYEPPKAKK 653
Db 181 FPQRFQDGRNDRYANRNVFFDINMLGLDGIQGPVYVGTGCVFNROALYDPPVSEKR 240
QY 654 -----PGFLASLCGGKKKASKSKRS-----SDKKK---SNKHVDSVVPFNL 693
Db 241 PKMTCDCWPSWCCCCCGSRKSKKKGKGLLGLLYGKKKMMKGNVYVKGSAPVFDL 300
QY 694 EDIEGVEGAGFDD-EKSVLMSQMSLEKRFQGSAAFAVASTLMEYGGVQSGSTPESLLKEA 752
Db 301 BEIEEGLE--GYEELEKSTLMSQNFKEKRFQSGSPVFIATIMENGGLPEGTNSTSIIKEA 358
QY 753 IHVISCYEDKSEWGTIEGNIYGSVTEIDILTGFKMHARGWRSVYCMKPAFKGSAPINL 812
```



Db 359 IHVISCYEEKTEKGWIGSVTIEDILTGFKMHCGRGWSVYCVKRPAPKGSAPINL 418  
Qy 813 SDRNLQVLRWALGSVEILFSRHCPLWYGYGRIKFLERFAYINTIYPLTSLPLLYCYIL 872  
Db 419 SDRHLQVLRWALGSVEILFSRHCPLWYGYGRIKFLERFAYINTIYPLTSLPLLYCYI 478  
Qy 873 PAICLLTGKFTMPETSINLASIFATGILEMRWSGVGIDWNRNEQFWIYGGI 932  
Db 479 PAVCLLTGKFIPLSNLTSVWFALFSLIATGVLELRWSGVSIQDWRNEQFWIYGGV 538  
Qy 933 SAHLFAVFOGLLKVLAGIDTNTVTSKANDDEGDFAEIYMFKWTLLIPIPTTILINMVG 992  
Db 539 SAHLFAVFOGLLKVLAGIDTNTVTSKANDDEGDFAEIYMFKWTLLIPIPTTILINMVG 597  
Qy 993 VVAGTSVAINSGYOSWGPLFKGLFFAFWVIVHLYPFLKGLMGRNQRTPTIIVLWVLLAS 1052  
Db 598 VVAGTSVAINSGYOSWGPLFKGLFFAFWVIVHLYPFLKGLMGRNQRTPTIIVLWVLLAS 657  
Qy 1053 IFSLWVRVDPFTTTLRAGPNITQTCGINC 1080  
Db 658 IFSLWVRVDPFTTTLRAGPNITQTCGINC 685  
RESULT 14  
US-09-838-539-7  
; Sequence 7, Application US/09838539  
; Patent No. US20020129401A1  
; GENERAL INFORMATION:  
; APPLICANT: Stalker, D. et al.  
; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter  
; FILE REFERENCE: 15621/03/US  
; CURRENT APPLICATION NUMBER: US/09/838,539  
; CURRENT FILING DATE: 2001-04-18  
; PRIOR FILING DATE: 1996-10-29  
; PRIOR FILING DATE: 1997-10-29  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 685  
; TYPE: PRT  
; ORGANISM: Gossypium hirsutum  
US-09-838-539-7  
Query Match 48.7% Score 2811; DB 10; Length 685;  
Best Local Similarity 74.1%; Pred. No. 2.9e-230; Indels 24; Gaps 6;  
Matches 509; Conservative 76; Mismatches 76;  
Qy 413 ARKWWPFVKYDIEPRAPEFYCQKIDYLDKQVPSFYKDRRAMKREYEEFKIRINALVS 472  
Db 1 ARRWVPFCCKHNVPRAPPEFYNEKIDYLDKQVPSFYKDRRAMKREYEEFKIRINALVA 60  
Qy 473 KALKVPEGWIMQDGTWPGNTRDHPGMIOVFLGHSGLDTEGNEPLRVYVSRKRP 532  
Db 61 KAQKKPEGWYMDGTWPGNTRDHPGMIOVFLGHSGLDTEGNEPLRVYVSRKRP 120  
Qy 533 FOHKKAGAMNALVRVSALVTNGOYMLNDCDHYINNSKAVREACMFMDPNLGPQVCYV 592  
Db 121 YQHHKKAGAEALVRVSALVTNAPFLNDCDHYINNSKAVREACMFMDPNLGPQVCYV 180  
Qy 593 QPQRFQDIDRDRYANRNTVFFDINLRLDGIQGVVYVGTGVFNRTAIYGYEPPKAK 652  
Db 181 QPQRFQDIDRDRYANRNTVFFDINLRLDGIQGVVYVGTGVFNRTAIYGYEPPKAK 240  
Qy 653 K-----PGFLASLGGKKKSKRS-----SDKKK---SNKHVDSSVPFN 692  
Db 241 RPKMTCDCWPSWCCCCGGSKKKKKKKGLGLGLGLGLGLGLGLGLGLGLGLGLGLGLGL 300  
Qy 693 LEDIEEGVEGAGFDD-EKSVLMSONSLKFRFQSAAFVASTIMEYGGVPOSSSTPESLLKE 751  
Db 181 LEDIEEGVEGAGFDD-EKSVLMSONSLKFRFQSAAFVASTIMEYGGVPOSSSTPESLLKE 751

Db 301 LEEIEGLE--GYEELEKSTLM  
Qy 752 AIHVISGYEDKSWGTEIGWI  
Db 359 AIHVISGYEEDKSWGTEIGWI  
Qy 812 LSDLNLQVLRWALGSVEILFSR  
Db 419 LSDLNLQVLRWALGSVEILFSR  
Qy 872 LPAICLLTGKFIPEISNLASI  
Db 479 IPAVCLLTGKFIPTLSNLTSTV  
Qy 932 ISAHLEAVFOGLLKVLAGIDTN  
Db 539 VSAHLFAVFOGLLKVLAGIDTN  
Qy 992 GVAGTSVAINSGYOSWGPLFG  
Db 598 GVAGTSVAINSGYOSWGPLFG  
Qy 1052 SIFSLWVRVDPFTTTLRAGPNIT  
Db 658 SIFSLWVRVDPFTTTLRAGPNIT  
RESULT 15  
US-09-900-237-20  
; Sequence 20, Application US/09  
; Patent No. US20020120124A1  
; GENERAL INFORMATION:  
; APPLICANT: Allen, Stephen  
; TITLE OF INVENTION: Plant Cell  
; FILE REFERENCE: BB1170 US CIP  
; CURRENT APPLICATION NUMBER: 0  
; CURRENT FILING DATE: 2001-07  
; PRIOR APPLICATION NUMBER: 60/  
; PRIOR FILING DATE: 1998-07-14  
; PRIOR APPLICATION NUMBER: PCT  
; PRIOR FILING DATE: 1999-07-13  
; PRIOR APPLICATION NUMBER: 09/  
; PRIOR FILING DATE: 2000-12-21  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: Microsoft Office 97  
; SEQ ID NO 20  
; LENGTH: 506  
; TYPE: PRT  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (88)  
US-09-900-237-20  
Query Match 46.5%  
Best Local Similarity 99.6%  
Matches 504; Conservative  
Qy 575 EAMCFMDPNLGPQVCYVQFPQ  
Db 1 EAMCFMDPNLGPQVCYVQFPQ  
Qy 635 CVFNRTAIYGYEPPKAKKPG  
Db 61 CVFNRTAIYGYEPPKAKKPG  
Qy 695 DIEEGVEGAGFDDKSVLMSON  
Db 121 DIEEGVEGAGFDDKSVLMSON  
Qy 755 VISCYEDKSWGTEIGWI  
Db 181 VISCYEDKSWGTEIGWI

EDILTGFKMHCGRGWSVYCVKRPAPKGSAPINL 811  
EDILTGFKMHCGRGWSVYCVKRPAPKGSAPINL 418  
YGYGRIKFLERFAYINTIYPLTSLPLLYCYI 871  
YGYGRIKFLERFAYINTIYPLTSLPLLYCYI 478  
ELSFATGILEMRWSGVGIDWNRNEQFWIYGG 931  
ELSFATGILEMRWSGVGIDWNRNEQFWIYGG 538  
KANDDEGDFAEIYMFKWTLLIPIPTTILINMV 991  
KANDDEGDFAEIYMFKWTLLIPIPTTILINMV 597  
FWIVHLYPFLKGLMGRNQRTPTIIVWVALLA 1051  
FWIVHLYPFLKGLMGRNQRTPTIIVWVALLA 657  
1078  
684  
e Syntheses  
900,237  
344  
9/15871  
33  
are 2689; DB 10; Length 506;  
ed. No. 4.2e-220;  
Mismatches 1; Indels 0; Gaps 0;  
IDRNDRYANRNTVFFDINLRLDGIQGVVYVGTG 634  
IDRNDRYANRNTVFFDINLRLDGIQGVVYVGTG 60  
EGGKKKSKSKRSDDKKSKSNKHVDSSVPFNLE 694  
EGGKKKSKSKRSDDKKSKSNKHVDSSVPFNLE 120  
RFGQSAAFVASTIMEYGGVPOSSSTPESLLKEAIH 754  
RFGQSAAFVASTIMEYGGVPOSSSTPESLLKEAIH 180  
ELTGFKMHCGRGWSVYCVKRPAPKGSAPINLUSD 814  
ELTGFKMHCGRGWSVYCVKRPAPKGSAPINLUSD 240

QY 815 RLNQVLRWALGSVEILFSRHCPLWYGGRLKFLERFAYINTTIYPLTSLPLLVYCILPA 874  
DB 241 RLNQVLRWALGSVEILFSRHCPLWYGGRLKFLERFAYINTTIYPLTSLPLLVYCILPA 300  
QY 875 ICLLTGKFIMPEISNLSIWFIALFSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISA 934  
DB 301 ICLLTGKFIMPEISNLSIWFIALFSIFATGILEMRWSGVGIDEMWRNEQFWVIGGISA 360  
QY 935 HLFVAVFQGLLKVLAGIDTNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTILIIINMGVV 994  
DB 361 HLFVAVFQGLLKVLAGIDTNTFTVTSKANDEEGDFAELYMFKWTTLLIPPTTILIIINMGVV 420  
QY 995 AGTSVAINSQSGWGLFGLFFAFWVIVHLYPFLKGLMGRONKPTTIVIVWAVLLASIF 1054  
DB 421 AGTSVAINSQSGWGLFGLFFAFWVIVHLYPFLKGLMGRONKPTTIVIVWAVLLASIF 480  
QY 1055 SLLWVRVDPFTTRLAGPNIOTCGINC 1080  
DB 481 SLLWVRVDPFTTRLAGPNIOTCGINC 506

Search completed: February 19, 2003, 16:41:44  
Job time : 39 secs

